

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 16:11:03 ; Search time 3009.5 Seconds

(without alignments)
10652.835 Million cell updates/sec

Title: SEQ1-69C

Perfect score: 564
Sequence: 1 gcagccgcgcctcagaaac.....agtgcgccttcgacttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_by:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562.4	99.7	10777	8 HS11Q13RP	Y12377 H. sapiens F
2	553	98.0	150110	14 AC097722	AC097722 Homo sapi
3	553	98.0	150214	8 AP006345	AP006345 Homo sapi
4	44.2	7.8	125020	8 AF429315	AF429315 Homo sapi
5	42.6	7.6	125020	8 AF429315	AF429315 Homo sapi
6	40.2	7.1	184533	8 AC093592	AC093592 Homo sapi
7	40.2	7.1	192573	8 AC073444	AC073444 Homo sapi
8	40.2	7.1	320658	14 AC145527	AC145527 Ateles
9	40	7.1	153129	9 AC163209	AC163209 Mus muscu
10	40	7.1	120041	9 AC111047	AC111047 Mus muscu
11	39.8	7.1	103034	14 AC067978	AC067978 Homo sapi
12	39.8	7.1	111934	8 AF281074	AF281074 Homo sapi
13	39.8	7.1	116039	8 AC007362	AC007362 Homo sapi
14	39.8	6.9	123561	14 AC159874	AC159874 Bos tauru
15	38.8	6.9	184532	9 AC124345	AC124345 Mus muscu
16	38.8	6.9	203106	9 AC021667	AC021667 Mus muscu
17	38.8	6.9	215604	14 AC160979	AC160979 Mus muscu
18	38.4	6.8	110000	14 CT005255_5	Continuation (6 of

C 19	38.2	6.8	2777	6 AX833454	AX833454 Sequence
C 20	38.2	6.8	2777	8 AK095212	AK095212 Homo sapi
C 21	38.2	6.8	74871	14 AC137095	AC137095 Homo sapi
C 22	38.2	6.8	178778	8 AC132936	AC132936 Homo sapi
C 23	38	6.7	110000	14 LMFCHR26_2	Continuation (3 of
C 24	38	6.7	119613	8 AC138582	AC138582 Pan trogl
C 25	38	6.7	121103	8 AP006623	AP006623 Homo sapi
C 26	38	6.7	142979	8 AC000029	AC000029 Homo sapi
C 27	38	6.7	190018	14 AC148588	AC148588 Pan trogl
C 28	38	6.7	205854	14 AC138186	AC138186 Homo sapi
C 29	37.8	6.7	111370	8 AC067815	AC067815 Homo sapi
C 30	37.8	6.7	134362	14 AC084853	AC084853 Homo sapi
C 31	37.8	6.7	138129	8 HS102D24	AL021391 Human DNA
C 32	37.6	6.7	114694	8 AL355500	AL355500 Human DNA
C 33	37.4	6.6	184866	9 AL611934	AL611934 Mouse DNA
C 34	37.4	6.6	187911	14 AC136867	AC136867 Rattus no
C 35	37.4	6.6	207735	9 AC130279	AC130279 Mus muscu
C 36	37.4	6.6	213590	14 AC154101	AC154101 Mus muscu
C 37	37.4	6.6	223357	14 AC111958	AC111958 Rattus no
C 38	37.4	6.6	240039	14 AC111926	AC111926 Rattus no
C 39	37.2	6.6	21670	8 AF508041	AF508041 Homo sapi
C 40	37.2	6.6	36628	8 AF037222	AF037222 Human DNA
C 41	37.2	6.6	294817	14 AC114626	AC114626 Mus muscu
C 42	37	6.6	1422	4 BT020874	BT020874 Bos tauru
C 43	37	6.6	1785	8 HSTPRMPL3	U68161 Human chrom
C 44	37	6.6	2500	1 SSU65940	U65940 Streptomyce
C 45	37	6.6	157585	5 BX322664	BX322664 Zebrafish

ALIGNMENTS

RESULT 1	HS11Q13RP	10777 bp	DNA	linear	PRI 02-MAR-2000
LOCUS	H. sapiens FGF-3 gene	upstream flanking region.			
DEFINITION	Y12377				
ACCESSION	Y12377.1	GI:1934871			
VERSION	Y12377.1	GI:1934871			
KEYWORDS	Alu repeat; FGF-3 gene; int-2 gene; L1 repeat; MIR repeat.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Dienabi,S., Brisson,O., Galdemand,C. and Lavialle,C.			
AUTHORS	1	Dienabi,S., Brisson,O., Galdemand,C. and Lavialle,C.			
TITLE	1	Sequence analysis of the transcription control region upstream of the human FGF-3 gene			
JOURNAL	1	DNA Seq. 10 (4-5), 317-329 (1999)			
PUBMED	10727086				
REFERENCE	2	(bases 1 to 10777)			
AUTHORS	Brisson,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (04-APR-1997) O. Brisson, Lab. de Genetique Oncologique, CNRSURA1967, Institut Gustave-Rousey, 39 Rue Camille Desmoulins, 94805 Villejuif CEDEX, FRANCE				
COMMENT	The BamHI site at the 3' end of this sequence (nt 10772) corresponds to the BamHI site at the 5' end of the FGF-3/int-2 gene sequence (nt 1, X1445), as ascertained by sequencing through this site.				
FEATURES	Related sequence: B04369.				
SOURCE	location/Qualifiers				
	1..10777				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="11"				
	/map="q13"				
	/tissue_type="placenta"				
	/clone_id="lambda D3, Clontech cat#HL1067J lot 1221"				
	/germline				
	489..631				
	/rpt_family="MIR2"				
	repeat_region				


```

misc_feature 4045..9798
/note="assembly_name:Contig26"
gap 9799..9898
/estimated_length=unknown
misc_feature 9899..14019
/note="assembly_name:Contig27"
gap 14020..14119
/estimated_length=unknown
misc_feature 14120..19392
/note="assembly_name:Contig28"
gap 19393..19492
/estimated_length=unknown
misc_feature 19493..24872
/note="assembly_name:Contig29"
gap 24873..24972
/estimated_length=unknown
misc_feature 24973..43358
/note="assembly_name:Contig30"
gap 43359..44058
/estimated_length=unknown
misc_feature 44059..68992
/note="assembly_name:Contig31"
gap 68993..69092
/estimated_length=unknown
misc_feature 69093..104669
/note="assembly_name:Contig32"
gap 104670..104769
/estimated_length=unknown
gap 104770..150110
/note="assembly_name:Contig33"
clone_end:SP6
vector_side:right"

```

```

ORIGIN
Query Match 98.0%; Score 553; DB 14; Length 150110;
Best Local Similarity 99.8%; Pred. No. 1.7e-121;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCGAAGAAAGAGAGCGACGACCTCAGGTGACTCACCCCATGTG 60
DB 50478 GCAGCCCTGCTCGAAGAAAGAGAGCGACGACCTCAGGTGACTCACCCCATGTG 50419

QY 61 GCTGAGGCGAGGAGGCTCTGAGGCGAGGCGAGGCGGCTGAGGTGATACCGCA 120
DB 50418 GCTGAGGCGAGGAGGCTCTGAGGCGAGGCGAGGCGGCTGAGGTGATACCGCA 50359

QY 121 GGGGTCTTGCCATGATGAGGCGACAGGGGCTGATACAGCTTACTAGTGACATCGAGTCC 180
DB 50358 GGGGTCTTGCCATGATGAGGCGACAGGGGCTGATACAGCTTACTAGTGACATCGAGTCC 50299

QY 181 CTGGTGCAGGCTCTGAAAGTCTGAAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 240
DB 50298 CTGGTGCAGGCTCTGAAAGTCTGAAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 50239

QY 241 GCCGGCCATGCCCCCAACGTTGCACTACCTGCTTTGAGAGGTTGGGGCTTCCAGTC 300
DB 50238 GCCGGCCATGCCCCCAACGTTGCACTACCTGCTTTGAGAGGTTGGGGCTTCCAGTC 50179

QY 301 ACAAGGTCCCATCCACGTAACCAAGCCCAAGGTGCTGCAAGAGTCCCTCGACGATGATAA 360
DB 50178 ACAAGGTCCCATCCACGTAACCAAGCCCAAGGTGCTGCAAGAGTCCCTCGACGATGATAA 50119

QY 361 CCAAGGAGGCTTTGGGAAACCACTGTAAGGCGATGGCTTTGATTTAAGTAGAGAGGTGG 420
DB 50118 CCAAGGAGGCTTTGGGAAACCACTGTAAGGCGATGGCTTTGATTTAAGTAGAGAGGTGG 50059

QY 421 GGGTGGGCTGGGCAAGGCGACCAAGTCTGAGTCAAGAGCCAGAGGCAAGAAAGTGTGCC 480
DB 50058 GGGTGGGCTGGGCAAGGCGACCAAGTCTGAGTCAAGAGCCAGAGGCAAGAAAGTGTGCC 49999

QY 481 A-GCAGTGCCTGGCGCTCTGCGATGAGTCTCTGAGCCACTGAGAAACAGCCTGTAGA 539

```

```

DB 49998 AGGACATGCCGCGCGCTCTGCGATGCACTCTCTGCCACCTGAGAAACGCTGTAGA 49939
QY 540 GAGGCACTGGCGCTTTTCGACATTC 564
DB 49938 GAGGCACTGGCGCTTTTCGACATTC 49914

```

RESULT 3

```

LOCUS AP006345/c 150214 bp DNA linear PRI 06-AUG-2005
DEFINITION Homo sapiens genomic DNA, chromosome 11 clone:RP11-186D19, complete
sequence.
ACCESSION AP006345
VERSION AP006345.4 GI:71891788
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

```

REFERENCE
AUTHORS Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,
Totohi,Y. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 150214)
AUTHORS Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,
Totohi,Y. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Aug 5, 2005 this sequence version replaced gi:48290861.

```

```

COMMENT
FEATURES
source
1..150214
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-186D19"

```

ORIGIN

```

Query Match 98.0%; Score 553; DB 8; Length 150214;
Best Local Similarity 99.8%; Pred. No. 1.7e-121;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCGAAGAAAGAGAGCGACGACCTCAGGTGACTCACCCCATGTG 60
DB 85889 GCAGCCCTGCTCGAAGAAAGAGAGCGACGACCTCAGGTGACTCACCCCATGTG 85830

QY 61 GCTGAGGCGAGGAGGCTCTGAGGCGAGGCGAGGCGGCTGAGGTGATACCGCA 120
DB 85829 GCTGAGGCGAGGAGGCTCTGAGGCGAGGCGAGGCGGCTGAGGTGATACCGCA 85770

QY 121 GGGGTCTTGCCATGATGAGGCGACAGGGGCTGATACAGCTTACTAGTGACATCGAGTCC 180
DB 85769 GGGGTCTTGCCATGATGAGGCGACAGGGGCTGATACAGCTTACTAGTGACATCGAGTCC 85710

QY 181 CTGGTGCAGGCTCTGAAAGTCTGAAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 240
DB 85709 CTGGTGCAGGCTCTGAAAGTCTGAAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 85650

QY 241 GCCGGCCATGCCCCCAACGTTGCACTACCTGCTTTGAGAGGTTGGGGCTTCCAGTC 300
DB 85649 GCCGGCCATGCCCCCAACGTTGCACTACCTGCTTTGAGAGGTTGGGGCTTCCAGTC 85590

QY 301 ACAAGGTCCCATCCACGTAACCAAGGCTGAGTCAAGAGTCCCTCGACGATGATAA 360
DB 85589 ACAAGGTCCCATCCACGTAACCAAGGCTGAGTCAAGAGTCCCTCGACGATGATAA 85530

```

QY 361 CCAAGGAGGCTTGGGAACCAATCTGAAGGGCATGCTTGTATTAGTAGAGAGGCTGG 420
Db 85529 CCAAGGAGGCTTGGGAACCAATCTGAAGGGCATGCTTGTATTAGTAGAGAGGCTGG 85470
QY 421 GCGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAAGGAGAGAACTGCTCCCC 480
Db 85469 GCGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAAGGAGAGAACTGCTCCCC 85410
QY 481 A-GCATGCGCCGCGCCTCTGCGATGACAGTCTCTGGGCACTGAGAAAGCCTGTAGA 539
Db 85409 AGGCATGCGCCGCGCCTCTGCGATGACAGTCTCTGGGCACTGAGAAAGCCTGTAGA 85350
QY 540 GAGCAGTGGCGCTCTTTCGCACTTC 564
Db 85349 GAGCAGTGGCGCTCTTTCGCACTTC 85325

RESULT 4
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315.1 GI:17646244
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
JOURNAL
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNAL
TITLE
location/Qualifiers
FEATURES
source 1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-like 2 (HDL2)"
Disease-like 2 (HDL2)"
complement(35581..35746)
/rpt_type=cadem
/rpt_unit="ctg"
complement(<36507..36887)
/gene="JPH3"
/note="synonym: JP3"
complement(<36507..36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSGGRFNFDDGSSYCGGMDKAGHGVCCTPKGQGEYTSWS
HGFEVLGYTPRSGNTYGTAAQGRHGIQLESKKWYKSGWITGFKGRIGVREACAG
NGAKYEGTWSNGLQDGYETETYSDC"

ORIGIN

Query Match 7.8%; Score 44.2; DB 8; Length 125020;
Best Local Similarity 12.0%; Pred. No. 3.2;
Matches 64; Conservative 222; Mismatches 244; Indels 2; Gaps 1;
QY 23 AAGACCGACGACACTACCGGTACTACCCCATGTGGCTGGAGGCGAGACCTCCT 82
Db 17541 MRSSKSSWSMSMRSSWCMGWGAGYRRSKRSASQMGWMSKKRSTGMRAQSKTGS 17482
QY 83 GAGCAGAGGCGGACGAGCGGTGAGTGGGAGCGAGAGGGGTCTTCCATGATGGGAC 142
Db 17481 YGSTRGSMKKKKYKSKRSRKGKKKCTCYCMKKYKTRKTSKMWYTKSGYRYRRC 17422
QY 143 AGGGCTGCATACATCTTACTAGTCAATGAGTCCCTGGTCCAGCCTTGAAGTC 202
Db 17421 MKKKGTGYRQMSKSKSGYSMRGMSYSTSCMSKWSYMWKCMKYSKKRBRMR 17362
QY 203 TGGAAATGAGCATATGTTCCCATTAAGAAAAGTGTGGCCGCGCATGCCCAAGCTT 262
Db 17361 GMSKSGMRGYAGRGYSSWMSSTRKRRSKCYSKYKKGKMKMGMRGSKYWSM 17302
QY 263 GCACACTCAGTGCCTTGGCAGAGGTGGGCTTCAGTCACAGGCTCCATCAGTACCA 322
Db 17301 KMKRSSWSKCTYKSKYSGRRSKGMRSTSKAKSGMRAGSKCTTGSSYMNRRNN 17242
QY 323 GCCCAGGTGGCTGCAAGAGT--CCCTGCAATGATGAACCAAGGAGCTTGGGAAAC 380
Db 17241 RMGTGCMYRRSRAMNNGNAAAAGCTTCCCATNMGGGGAAAAGGCGSASBASCY 17182
QY 381 CACATCTGAAGGCGATGCTTATTTAGTGAAGAGGTGGGCTGGCTGGGCAAGGCCA 440
Db 17181 KGRMSKSCYRSYRRRCMSKCRGSRGSMGTRRSQKTSYSASGSRGCTYCGW 17122
QY 441 CCAGTCTGAGTCTGAGCAGGACGAGGAGGAGAGCTGTCCCGCAGCTGCCCGCCTCTG 500
Db 17121 GRKRKCMKSKMKMYTSYTRKRKMTCKMKCYMTAAMYRSKCMCMCSGCCGCTYMG 17062
QY 501 CGATGCACTCTCTCTGCGACCTGAGAACAGCTGTAGAGAGCGTGGCGT 552
Db 17061 MSSYSYGKYSWGMKSYMMRSYYSKRSTSKAMRSSKMGTGGRYKGGGRSY 17010

RESULT 5
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315.1 GI:17646244
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
JOURNAL
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNAL
TITLE
location/Qualifiers
FEATURES
source 1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"

	repeat_region	/map=16q24.3; between DIS520 and WI-12410" /note="isolated from a patient with Huntington's Disease-Like 2 (HDL2)" complement(3581..35746) /rpt_type=tandem /rpt_unit="ctg" complement(<36507..>36887) /gene="JPH3"
gene		/note="synonym: JP3" complement(<36507..>36887) /gene="JPH3" /product="junctionillin 3" complement(<36507..>36887) /gene="JPH3"
mRNA		/note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctionophilin 3" /protein_id="P44094.1" /db_xref="GI:17646245" /translation="MSGGFNFNDGGSYTCGGNEDKAHGHCCTGPKGGEYTSGWS HGFELGVATWPSGNTYQGTGAQAQRKHGIGLESKGMVYKGEWTGPKRGVGRACAG NGAYEGEWTSGNLQDGYETRESDG"
CDs		
ORIGIN		
Query Match	7.6%; Score 42.6;	DB 8; Length 125020;
Beeb Local Similarity	11.5%; Pred. No. 7.6;	
Matches	45; Conservative 170; Mismatches 175; Indels 0; Gaps 0;	
Oy	32 GCACACCAACGCGTACTCACCCCATGTGCTGGAAGCCAGGAGCCTCTGAGCAGG 91	
Db	51746 KSAISCMTWBSKACKSSYCAKRSGLCYKKTWGSMYTSGRSMMYTTSTYSTCMTHMWMWSY 51805	
Oy	92 CCAGGCAAGCCGTCAGCTGGGTGCAGCGCAGGCGCTTGCCATGCTGGACAAGGCGCTTC 151	
Db	51806 MSWYSKCYRTYSCKRKNKCWCMSGYSRBACCTWRSMWTSCTWYSRBTGMNRKMGWCWCY 51855	
Oy	152 ATACAGCTTACTCAGTACATCGAAGTCCCTGGTCCAGCCTCTGGAAGTCTGGAAGTGA 211	
Db	51866 RRCGCMGKWGYCKGKKSWSMKCATTTGMSCAKRSMSVCGRKCRPASCMGCYKCMSMT 51925	
Oy	212 GCATGCTTCCCATTAAGAAAGTGCTGCCGCCGCATGCCCCCCCAAGCTTGACACTCA 271	
Db	51926 CMMKSYTSCCTCYGACYCMMKGRKRRSSYKSCYKKSMTKXSMGMSCYGRSRMYSY 51985	
Oy	272 CTGCTTTGAGAGGTTGGGGCTTCCAGTCAAGAGTCCCATCCATCCAGTACCAGCCAGGTG 331	
Db	51986 RYRKRKNCMWMTTRKSGKGSYNSSWCMGAKGTSTMCASSYAMSCKRMGSATSNGCC 52045	
Oy	332 GTTCGAGAAGGTCCCTCGCAGTCATGAACCAAAGGAGCTTGGAAACCAATCTGAAG 391	
Db	52046 YMSYRSASMCWGSTGSRYSKYSTRSCWSCMRGSMYSRBMKGAGGCACMSCYRKSRK 52105	
Oy	392 GGCAATGGCTTTGATTGATGAGAGCGGTGG 421	
Db	52106 RGSWKGSMTKSKXTGKGMWGRGSGWSKS 52135	
RESULT 6		
LOCUS	AC093592/c	184533 bp DNA linear PRI 13-MAY-2005
DEFINITION	Homo sapiens BAC clone Rp11-50L21 from 4, complete sequence.	
ACCESSION	AC093592	
VERSION	AC093592.3	GI:16756323
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 184533)	Meyer,R., Doeberer,A. and Bielicki,L.
AUTHORS		The sequence of Homo sapiens BAC clone Rp11-50L21
TITLE		

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished (2001)
2 (bases 1 to 184533)
Waterston,R.H.
Direct Submission
Submitted (05-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 184533)
Waterston,R.H.
Direct Submission
Submitted (07-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

4 (bases 1 to 184533)
Waterston,R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

5 (bases 1 to 184533)
Waterston,R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
(bases 1 to 184533)
Wilson,R.K.
Direct Submission
Submitted (13-MAY-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT

On Nov 7, 2001 this sequence version replaced gi:15687339.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0050121

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Cateneese,V.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at <http://www.chori.org>

VECTOR: pRACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-628B17. Actual start of
this clone is at base position 1 of RP11-50L21; actual end is at
base position 184533 of RP11-50L21.

Data from AC073344 and AC016032 was used to finish this clone.
AC093592. Polymorphisms have been identified between AC073344 and

```

FEATURES             AC093592.
source                1..184533
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="RP11-50L21"
                     /clone_1fb="RPC1-11"
                     /complement(19506..26531)
                     /gene="KIAA1430"
                     /complement(join(19506..20820,26266..26531))
                     /gene="KIAA1430"
                     /complement(join(20779..20820,26266..26531))
                     /gene="KIAA1430"
                     /complement(join(20779..20820,26266..26531))
                     /note="Homo sapiens KIAA1430 protein, mRNA (CDNA clone
                     MGC:40366 IMAGE:5240222), complete cds.; H_NH0050121.1
                     This gene was based on gi(20988136)
                     Continued from H_NH0628B17.1"
                     /codon_start=3
                     /product="unknown"
                     /protein_id="AAV40976.1"
                     /db_xref="gi:6392942"
                     /translation="FLQIDKGPQNHFDQSVADGKNYSFTRESVRQIDENQRLK
                     ELNRQAEKRGSKSTIPRSADHPKLYHSALNRQKQRIERENLTISLCHGWSAVT"
                     54035..54810
                     /note="CpG_island (%GC=74.2, o/e=0.97, #CpGs=110)"
                     59958..61769
                     /note="CpG_island (%GC=76.4, o/e=0.85, #CpGs=171)"
                     72793..73003
                     /note="CpG_island (%GC=58.8, o/e=1.06, #CpGs=17)"
                     90414..91446
                     /note="CpG_island (%GC=71.1, o/e=0.68, #CpGs=80)"

ORIGIN
Query Match          7.1%; Score 40.2; DB 8; Length 184533;
Best Local Similarity 60.6%; Pred. No. 27;
Matches 66; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY      18 AACAGAGGAGCGACACACTCAGCGTGACTGACCCCATGTGCGAGCGGAGGAGGC 77
Db      60170 AGCTCCCGGCGCGGCGGCGGCGGAGGCCCACTGCTGTGTGACTGAAGCTTGCGCCAGC 60111

QY      78 CTCCTGAGGCGAGCGCGCAGCGGCGGCGTCAGTGAGTGACGCGCAGCGGCGTC 126
Db      60110 CTGAGCGCGGCGGCGGCGGCGCGCCCTTAAGTGCCTGTGCGGCGGCGGC 60062

RESULT 7
AC073344/c          192573 bp   DNA       linear   PRI 16-APR-2002
LOCUS              AC073344
DEFINITION         Homo sapiens BAC clone RP11-628B17 from 4, complete sequence.
ACCESSION          AC073344
VERSION            AC073344.7 GI:19848493
KEYWORDS           HTG.
SOURCE             Homo sapiens (human)
ORGANISM           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 192573)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
JOURNAL            Genome Res. 8 (11), 1097-1108 (1998)
PUBMED            9847074
2 (bases 1 to 192573)
Harris,A., Meyer,R. and Nguyen,C.
The sequence of Homo sapiens BAC clone RP11-628B17
JOURNAL            Unpublished (2001)
3 (bases 1 to 192573)
Waterston,R.H.
Direct Submission
JOURNAL            Submitted (14-JUN-2000) Genome Sequencing Center, Washington

```

```

REFERENCE
AUTHORS           Watserson,R.H.
TITLE             Direct Submission
JOURNAL           Submitted (15-FEB-2002) Genome Sequencing Center, Washington
Univrsity School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 192573)
Waterston,R.H.
Direct Submission
Submitted (30-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 192573)
Waterston,R.
Direct Submission
Submitted (16-APR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 30, 2002 this sequence version replaced gi:18677646.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watserson.wustl.edu
----- Summary Statistics
Center project name: H_NH0628B17
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenoe,M., Catanesse,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION: The clone sequenced to the right is RP11-714G18. Actual start of this clone is at base position 1 of RP11-628B17; actual end is at base position 192573 of RP11-628B17.

Polymorphisms exist between AC073344 and AC093592. Data from AC093592 was used to finish AC073344.

```

FEATURES             AC093592.
source                1..192573
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="4"
                     /clone="RP11-628B17"

```

```

repeat_region
381..745
/rpt_family="MER1_cype"
misc_feature
742..1072
/note="match to EST BF244578 (NID:g11158509) "
742..856
/note="match to EST AA164371 (NID:g17411318) zp04a09.r1"
857..1424
/note="similar to EST BM511659 (NID:g18682802) tjs1d11.x1"
922..1218
/note="similar to Homo sapiens EST H84319 (NID:g1062990)
y897a10.r1"
misc_feature
1003..1811
/note="match to EST BG61606 (NID:g13667377) "
1079..1331
/note="match to EST AA309603 (NID:g1961983) "
1228..1583
/note="match to EST H84320 (NID:g1062991) y897a11.r1"
1251..1286
/rpt_family="GA-rich"
misc_feature
1255..1811
/note="match to EST BF754577 (NID:g15746155) "
1367..1811
/note="match to EST BF114916 (NID:g10984392) "
1383..1731
/note="similar to Mus musculus EST BB651105
(NID:g16485359) "
1466..1811
/note="similar to EST BM729997 (NID:g19051330) tjs1d11.y1"
1526..1784
/note="match to EST AA384441 (NID:g2036761) "
1628..1780
/note="match to EST BE083868 (NID:g8474218) "
1670..1815
/note="similar to Homo sapiens EST BF172452
(NID:g13438666) "
2149..2457
/rpt_family="Alu"
2255..2256
/note="match to EST H84320 (NID:g1062991) y897a11.r1"
3136..3164
/rpt_family="AT-rich"
3330..3634
/rpt_family="Alu"
3635..3919
/rpt_family="L1"
3965..4273
/rpt_family="Alu"
4375..4510
/rpt_family="Alu"
4590..4803
/rpt_family="MER1_cype"
4806..5101
/rpt_family="Alu"
5136..5290
/rpt_family="Alu"
5530..5609
/rpt_family="MR"
5610..5726
/rpt_family="L1"
5811..6097
/rpt_family="Alu"
6519..6552
/rpt_family="tRNA-Gln-CAG"
6679..6790
/rpt_family="MER1_cype"
6693..6698
/note="match to EST AW881803 (NID:g8043813) "
6791..7081
/rpt_family="Alu"
7082..7219
/rpt_family="MER1_cype"
7417..7722

```

[illegible]

Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorrie, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B., Hall, J., Horton, L., Hulne, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Piere, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teefaye, S., Theodore, U., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viet, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Journal
Submitted (07-JUN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
Britten, B., Nusbbaum, C., Lander, E., Abouelell, A., Allen, N., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorrie, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B., Hall, J., Horton, L., Hulne, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Piere, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teefaye, S., Theodore, U., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viet, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Journal
Submitted (03-AUG-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Aug 3, 2005 this sequence version replaced gi:67003668.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES
source
1. 153129
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP24-11708"
/clone_lib="RPICT-24 Male Mouse BAC"
2843. .2893
/rpt_family="AT_r1ch"
repeat_region
complement(2985. .3180)

Center: Broad Institute of MIT and Harvard
Center code: W1BR
Web site: http://www-seq.w1.mit.edu
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: 153129
Center clone name: 117_O_8

repeat_region
5441. .5470
/rpt_family="B4"
repeat_region
5748. .5876
/rpt_family="GAA)n"
repeat_region
6984. .7110
/rpt_family="TCTA)n"
repeat_region
7462. .7801
/rpt_family="MLTIC"
repeat_region
complement(9528. .9910)
/rpt_family="ORR1D"
repeat_region
9937. .10085
/rpt_family="MT2B"
repeat_region
10825. .10982
/rpt_family="MER5B"
repeat_region
complement(11051. .11261)
/rpt_family="TG)n"
repeat_region
12012. .12033
/rpt_family="B4"
repeat_region
12907. .13123
/rpt_family="A)n"
repeat_region
13433. .13812
/rpt_family="MLTIC"
repeat_region
complement(14278. .14503)
/rpt_family="WTE"
repeat_region
complement(14517. .14746)
/rpt_family="L1"
repeat_region
14745. .15433
/rpt_family="L1_MM"
repeat_region
complement(15710. .15820)
/rpt_family="L1MA4A"
repeat_region
complement(15823. .16050)
/rpt_family="L1MA4A"
repeat_region
16057. .16122
/rpt_family="TG)n"
repeat_region
complement(16123. .16262)
/rpt_family="B3A"
repeat_region
16576. .16671
/rpt_family="TG)n"
repeat_region
16676. .16856
/rpt_family="TG)n"
repeat_region
16862. .16925
/rpt_family="TATATG)n"
repeat_region
16928. .17091
/rpt_family="TATG)n"
repeat_region
complement(17311. .17610)
/rpt_family="MLT1A0"
repeat_region
complement(17617. .19453)
/rpt_family="L1_MM"
18133. .18163
/note="single clone coverage"
19454. .20704
/rpt_family="L1"
repeat_region
20758. .20824
/rpt_family="MIR"
repeat_region
complement(20852. .21030)
/rpt_family="MIR"
21915. .21942
/rpt_family="TTTTTG)n"
repeat_region
complement(21943. .22157)
/rpt_family="B3"
repeat_region
22301. .22397
/rpt_family="L2"
repeat_region
complement(22732. .22787)
/rpt_family="L1"
repeat_region
22786. .23404
/rpt_family="L1_MM"
repeat_region
complement(23405. .23879)
/rpt_family="MERVL_LTR"
23880. .24023
/rpt_family="L1_MM"
repeat_region
24051. .24205
/rpt_family="TCTCC)n"

```

repeat_region      25614..25638
                    /rpt_family="AT_rich"
repeat_region      26685..26707
                    /rpt_family="(TTTG)n"
repeat_region      28871..28954
                    /rpt_family="(TA)n"
repeat_region      complement(29240..29627)
                    /rpt_family="TAPEY_LTR"
repeat_region      complement(29660..29993)
                    /rpt_family="ORR1D"
repeat_region      30197..30223
                    /rpt_family="(TTTC)n"
repeat_region      complement(30226..30374)
                    /rpt_family="B1_MM"
repeat_region      30733..30785
                    /rpt_family="(TTTC)n"
repeat_region      complement(30891..31095)
                    /rpt_family="MTR"
unSURE             complement(30926..31037)
                    /note="single clone coverage"
repeat_region      3116..31151
                    /rpt_family="(GGGGA)n"
unSURE             31130..31158
                    /note="single clone coverage"
repeat_region      31345..31410
                    /rpt_family="(TG)n"
repeat_region      complement(31417..31571)
                    /rpt_family="RSINEL"
repeat_region      complement(3156..33865)
                    /rpt_family="RMR6A"
repeat_region      38071..38261
                    /rpt_family="B3A"
repeat_region      38280..38428
                    /rpt_family="B3A"
repeat_region      38402..38558
                    /rpt_family="(CA)n"
repeat_region      39152..39203
                    /rpt_family="(TTA)n"

```

Query Match Best Local Similarity 7.1%; Score 40; DB 9; Length 153129;

Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

QY 301 ACAGGGTCCATCACTACAGTACAGCCAGGCTGTCGAGAGAGTCCCTCCAGTCATGAAA 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76180 ACAGGATACCTCTGTGGAGCCAGGCGAGAGGCGAGGCGCCAGTCAGAGGCAAG 76121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 CCAAGGAGGCTTTGGAAACACATCTGAAGGCGATGCTTTGATTATTAGTGAGGGTGG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76120 GACAGGAGCCAGGCGGTGGAGCTAGGAGGAGGCGGCGGCCAGAGCAGAGGCGAGG 76061
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 GGCTGGGCTGGGAGGCCACAGGCTGTAGTCAGAGCCAGAGGCGAGG 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76060 GGCGGGGCGAGGCGAGGAGGCGAGGCGAGGCGAGGCGAGGCGAGG 76013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10 AC111047 230041 bp DNA linear ROD 10-OCT-2003

DEFINITION Mus musculus chromosome 3, clone RP23-176H24, complete sequence.

AC111047

VERSION AC111047.9 GI:37620322

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 230041)

Authors Birren, B., Nusbaum, C. and Lander, E.

Journal Mus musculus chromosome 3, clone RP23-176H24

REFERENCE Unpublished

2 (bases 1 to 230041)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhagaler, B., Brown, A., Camarata, A., Campopiano, A., Chang, J., Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kander, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Lander, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., Meldrum, J., Menus, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

FEATURES

source

Submitted (10-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 10, 2003 this sequence version replaced gi:34495134.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L19822
Center clone name: 176_H_24

----- Location/Qualifiers
1..230041
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"

JOURNAL
REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 103034)
Birtten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Balding, J., Barna, N., Bastien, V., Bede, F.,
Boguslavsky, L., Boutin-Galter, B., Brown, A., Burkett, G.,
Campobiano, A., Castele, A., Choepell, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karacas, A.,
Klein, J., Laroque, K., Lamasaras, R., Landers, T., Lenoczky, J.,
Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheeters, R.,
Melatim, J., Meneus, L., Mihova, T., Miranda, C., Mianga, Y., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange, Thoman, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessie, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 3, 2000 this sequence version replaced gi:7656720.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L9518
Center clone name: 2350 X 7
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 95545 bases at least Q40
Consensus quality: 99446 bases at least Q30
Consensus quality: 100858 bases at least Q20
Insert size: 98000; agarose-fp
Insert size: 101634; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1585: contig of 1585 bp in length
* 1586 1585: gap of 100 bp
* 1586 3150: contig of 1465 bp in length
* 3151 3250: gap of 100 bp
* 3251 4742: contig of 1492 bp in length
* 4743 4842: gap of 100 bp
* 4843 6347: contig of 1505 bp in length
* 6348 6447: gap of 100 bp
* 6448 8791: contig of 2343 bp in length
* 8791 8891: gap of 100 bp
* 8891 11684: contig of 2794 bp in length
* 11685 11785: gap of 100 bp
* 11785 15905: contig of 4121 bp in length
* 15906 16006: gap of 100 bp
* 16006 21395: contig of 5390 bp in length

FEATURES
source
* 21396 21495: gap of 100 bp
* 21496 26064: contig of 4569 bp in length
* 26065 26164: gap of 100 bp
* 26165 34912: contig of 8748 bp in length
* 34913 35012: gap of 100 bp
* 35013 46466: contig of 11454 bp in length
* 46467 46566: gap of 100 bp
* 46567 56982: contig of 10416 bp in length
* 56983 57082: gap of 100 bp
* 57083 71308: contig of 14226 bp in length
* 71309 71408: gap of 100 bp
* 71409 86559: contig of 15151 bp in length
* 86560 86659: gap of 100 bp
* 86660 103034: contig of 16375 bp in length.
Location/Qualifiers
1..103034
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="CTD-2350K7"
/clone_lib="CTD Human BAC"
1..1585
/note="assembly_fragment"
1586..1685
/estimated_length=100
1686..3150
/note="assembly_fragment"
3151..3250
/estimated_length=100
3251..4742
/note="assembly_fragment"
4743..4842
/estimated_length=100
4843..6347
/note="assembly_fragment"
6348..6447
/estimated_length=100
6448..8790
/note="assembly_fragment"
8791..8890
/estimated_length=100
8891..11684
/note="assembly_fragment"
11685..11784
/estimated_length=100
11785..15905
/note="assembly_fragment"
15906..16005
/estimated_length=100
16006..21395
/note="assembly_fragment"
21396..21495
/estimated_length=100
21496..26064
/note="assembly_fragment"
26065..26164
/estimated_length=100
26165..34912
/note="assembly_fragment"
34913..35012
/estimated_length=100
35013..46466
/note="assembly_fragment"
46467..46566
/estimated_length=100
46567..56982
/note="assembly_fragment"
56983..57082
/estimated_length=100
57083..71308
/note="assembly_fragment"

ORIGIN	Query Match	7.1%; Score 39.8; DB 14; Length 103034;
Query Match	Beat Local Similarity 64.8%; Pred. No. 36;	
Matches 59; Conservative	0; Mismatches 32; Indels 0; Gaps 0;	
gap	86560.86659	
mlac_feature	/estimated length=100	
	86660.103034	
	/note="assembly_fragment"	
Db	93047 TGATGAGAGGTGTTCCCAAGGCGCTGGCCA 93077	
Oy	64 GGAGCGGAGGAGCGCTCTGAGCGAGGCCCA 94	
Oy	4 GCCTGCTCTGAAAACAGAGAGCGACGACACTACGGTACTACCCCATGTGCT 63	
Db	92987 GCCCTCCCTCGGAAACGAGAGAGCTGGGGCCTCCGCGGGCTCACCTCATTTGGAA 93046	
RESULT 12		
AF281074/c		
LOCUS	AP281074 111934 bp DNA linear PRI 21-DEC-2000	
DEFINITION	Homo sapiens neuropilin 2 (NRP2) gene, complete cds, alternatively spliced.	
ACCESSION	AF281074	
VERSION	AF281074.1 GI:11934947	
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo;	
REFERENCE	1 (bases 1 to 111934)	
AUTHORS	Rossignol, M., Gagnon, M.L. and Klasebrun, M.	
TITLE	Genomic organization of human neuropilin-1 and neuropilin-2 genes: identification and distribution of splice variants and soluble isoforms	
JOURNAL	Genomics 70 (2), 211-222 (2000)	
REFERENCE	2 (bases 1 to 111934)	
AUTHORS	Rossignol, M., Gagnon, M.L. and Klasebrun, M.	
TITLE	Direct Submission	
JOURNAL	Submitted (21-JUN-2000) Surgical Research, Children's Hospital, 320 Longwood Ave, Boston, MA 02115, USA	
FEATURES	Location/Qualifiers	
source	1. 111934	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	<3518..>111422	
gene	/gene="NRP2"	
mrna	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63142..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87023,108610..108660,111118..>111422)	
	/gene="NRP2"	
	/product="neuropilin-2a(22)"	
	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63142..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87008,108610..108660,111118..>111422)	
	/gene="NRP2"	
	/product="neuropilin-2a(17)"	
	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63142..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87023,108610..108660,83879..84141,85679..85775,86988..87023,108610..108660,	
CDS		

```

11118. 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2a (22)"
/protein_id="AA641898.1"
/db_xref="GI:11934949"
/translation="MDMPFLTWVFLALYRSRQVRCQDDPCGGRINSKAGYITSPG
YLPQYSPHONCEWIAVPAPEPNQKIVLNFNPHFEIRKQCKDFEIRRGDSBSDLIG
KHCNINAPRITISSGMLYIKFTPSDYARQAGAFSLRYEIPFTSGDSKFNSTENGT
ESPPPEPKYPHNLDCFTFLIAKRMELIIQPLIFLDELHPQVSGDCQYMLIMQI
IPHGLPIKRCGCTKPSSELRSTGLSLTFHTMDAVAKDGSAAVYVYVHPEPLENO
CNVPLGMESSRIANEQISASTYSDSRWTPOOSRLHGDDNGMTPLDSNKELYQVDR
FLTMILALIQGALSRFTQNGYVKSYLEVSTNEDMVTYRHGNCVHPFANDDATE
VLNKLHAPLIRPVRIRPQTWHSIALLEIFGCRVTDAPCSMLNGLSLINDSOI
SASTOEWLSPSARLVSSRSRGMPFRIQAPQGEMLQVLDGTYKTVGVIICGAG
GDSITAEARAFPRKFKVSYSLNGKMEYIOPRPQOPLFEGMNHVDTPIRPFDI
PAQVRYPRMSPAGIGRLEVLGCDMTDSRPTVETGLPVKSEETTPPYTEEBAT
BRCNGSFEEDDKQILPSPGNCFDELTPQSNMDDAKMRLRTWMASSSPRSDPDS
DNRVLRLQSSQREGOYARLISPPHLPSPYCMWFQYQATGRCGVALQVYREASQS
KLWLVIREDDGSQWKHGRILIPSYDMEYQIVEGVYIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENEKVDIPEIHREBGEYDEIDEXYVDMSSNSATSGSGASPSTDEK
SMVLTLPIRITIMASGLVLTGATCAGALLCYTCYSYGLSRSCITLLENVPELYD
GLKKHVMNHQKCSBA"
join(3518. 3590,17760. 17937,36403. 36584,42688. 42918,
43995. 44150,46123. 46292,48101. 48256,60728. 60872,
63412. 63761,65955. 66099,69934. 70050,73043. 73183,
83879. 84141,85679. 85775,86988. 87008,108610. 108660,
11118. 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2a(17)"
/protein_id="AA641897.1"
/db_xref="GI:11934948"
/translation="MDMPFLTWVFLALYRSRQVRCQDDPCGGRINSKAGYITSPG
YLPQYSPHONCEWIAVPAPEPNQKIVLNFNPHFEIRKQCKDFEIRRGDSBSDLIG
KHCNINAPRITISSGMLYIKFTPSDYARQAGAFSLRYEIPFTSGDSKFNSTENGT
ESPPPEPKYPHNLDCFTFLIAKRMELIIQPLIFLDELHPQVSGDCQYMLIMQI
IPHGLPIKRCGCTKPSSELRSTGLSLTFHTMDAVAKDGSAAVYVYVHPEPLENO
CNVPLGMESSRIANEQISASTYSDSRWTPOOSRLHGDDNGMTPLDSNKELYQVDR
FLTMILALIQGALSRFTQNGYVKSYLEVSTNEDMVTYRHGNCVHPFANDDATE
VLNKLHAPLIRPVRIRPQTWHSIALLEIFGCRVTDAPCSMLNGLSLINDSOI
SASTOEWLSPSARLVSSRSRGMPFRIQAPQGEMLQVLDGTYKTVGVIICGAG
GDSITAEARAFPRKFKVSYSLNGKMEYIOPRPQOPLFEGMNHVDTPIRPFDI
PAQVRYPRMSPAGIGRLEVLGCDMTDSRPTVETGLPVKSEETTPPYTEEBAT
BRCNGSFEEDDKQILPSPGNCFDELTPQSNMDDAKMRLRTWMASSSPRSDPDS
DNRVLRLQSSQREGOYARLISPPHLPSPYCMWFQYQATGRCGVALQVYREASQS
KLWLVIREDDGSQWKHGRILIPSYDMEYQIVEGVYIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENEKVDIPEIHREBGEYDEIDEXYVDMSSNSATSGSGASPSTDEK
SMVLTLPIRITIMASGLVLTGATCAGALLCYTCYSYGLSRSCITLLENVPELYD
GLKKHVMNHQKCSBA"
join(3518. 3590,17760. 17937,36403. 36584,42688. 42918,
43995. 44150,46123. 46292,48101. 48256,60728. 60872,
63412. 63761,65955. 66099,69934. 70050,73043. 73183,
83879. 84141,85679. 85775,86988. 87008,108610. 108660,
11118. 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2b(5)"
/protein_id="AA641900.1"

```


overlap; the clone sequenced to the right is RP11-394E1, 200 base pair overlap. Actual start of this clone is at base position 117994 of RP11-335M10; actual end is at base position 115845 of RP11-150F11.

FEATURES

Source

Location/Qualifiers

1..116039

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/clone="RP11-150F11"

/clone_11b="RP11-11"

12630..13592

/note="CpG island (%GC=70.3, o/e=0.84, #CpGs=108)"

14041..97553

/gene="NRP2"

/gene="NRP2"

join(14041..14113,28294..28471,46943..47124,53228..53458,

54535..54690,56663..56832,58641..58796,71269..71413,

73953..74302,76496..76640,80475..80591,83585..83725,

94424..94686,96224..96320,97533..97553)

/gene="NRP2"

/note="Homo sapiens neuropilin 2 (NRP2), mRNA;"

H_NH0150F11.1

This gene was based on gi(4505458)

Continues as H_NH0394E01.1"

/codon_start=1

/product="unknown"

/protein_id="AA93216.1"

/db_xref="gi:62702291"

/translation="MDNFPPLTWFLALYFSNRHQVRCQDPPECGRLNSKDAQYTSRG

YPOYPSHONCEWIVIAEPNOQIVLNFPHFELIKDCKDPLRIDGSESDLLG

KHCNIAPIPTISSGSMYIRFTSDYVAFQGFSLRYEIFKTSGEDSKPTSPNGIT

ESPPPEKYPHNLCTFTLLAKPKMEIILQPLIDLEHDPLOVGEQCKDMLIMQ

IPHYPLIGKYGCTKPTSELSSGILSLTFPTDMAVAKGFSARYVIAHQEPLENO

CNVPLGMESGRIANEQISASTSYSDKRTPOSLHGDNDKMTNLSNKEYLOYDIR

FLTMLTATGASISRETONGYVYKSYKLEVTNGEEMVVRHKNKHPQANDDATE

VNLKLNAPLTLTRVRIAPQTHSGIALRELFCRCVYDAFCNSMLKMLISLADSOI

SASSTOELTMSPSAARLVSSRSQWFPRIPOAQPESEMLQVLDPTKTVKGIIDGAG

GDSITAEARAFVRKFKYSISLNGKMEYIDDPRTPOOLKEEGMHYDTPDIRFEDI

PAOVYRVPERMSPAGIGMRLVLCGDMTSCPTVETLIGPVYKEETTPPEEERT

ECGNCSEFEDDKDQLPDSGRFCNDFLEBPCGMVYDAKMLRTTMASSSSPNDTFPD

DRNFLRQSDQREGQYARLISPPVHLPRSVCMFPYQATGSGVALQVRENSQES

KLMVIREDDQGEWKGRIILPSYDMEQYVFEVIGKSGRGEIADDIRISTDVPLE

NMEPISAF"

15626..17342

/gene="NRP2"

/note="CpG island (%GC=64.9, o/e=0.68, #CpGs=145)"

ORIGIN

Query Match

Best Local Similarity 7.1%; Score 39.8; DB 8; Length 116039;

Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Db 77259 GGCCTGCTTCAGAAAAGAGAGACGACACTCAAGTACTCAACCCCATGTGCT 63

GACCTCCCTCGGAACAGAGAGCTGGGACACCTCGTGCTCAACCTCATTTGGAA 77200

GAGAGCGAGGAGCGCTCTGAGGAGCGCCA 94

Db 77199 TGATGAGAGGTGTTCCCAAGGCTGGGCCA 77169

RESULT 14

AC159874/c

LOCUS AC159874 233561 bp DNA linear HTG 01-JUL-2005

DEFINITION Bos taurus clone CH240-5418, *** SEQUENCING IN PROGRESS ***, 26

ACCESSION

AC159874

VERSION

AC159874.3 GI:68267115

KEYWORDS

HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE

Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

Munry D,Marie, Metzker M, Lee, Abramson S, Adams C, Alder J,

Allen C, Allen H, Albrooke S, Amin A, Angiano D,

Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H,

Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F,

Blewalo K, Blair D, Blankenburg K, Blyth P, Brown M,

Bryant N, Buhay C, Burch P, Burrell K, Calderon A,

Cardenas V, Carter K, Cavazos I, Caesar H, Center A,

Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Chu J,

Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L,

Davila M, L, Davis C, Davy-Carroll L, De Anda C, Dederich D,

Delgado O, Denon S, Deramo C, Ding Y, Din H, Divya K,

Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K,

Egan A, Escotto M, Eugene C, Evans C, A, Falle T, Fan G,

Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P,

Fraser C, M, Gabisi A, Ganta R, Garcia A, Garner T, Garza M,

Gebregeorgis E, Geer K, Gill R, Grady M, Guerra M, Guevara W,

Gunnarstne P, Healand W, Hamli C, Hamilton C, Hamilton K,

Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J,

Hernandez R, Hines S, Hladun S, L, Hodgson A, Hognes M,

Hollins B, Howell S, Huliy S, Hume J, Idebird D, Jackson A,

Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A,

Karpachy S, Kelly S, Kelly S, Khan Z, King L, Kovar C,

Kowis C, Kraft C, L, Lebow H, Levan J, Lewis L, Li Z, Liu J,

Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J,

Lorenshuwa L, Loulsegged H, Lozada R, J, Lu X, Ma J,

Maheshwari M, Mahindartne M, Mahmoud M, Malloy K, Mangum A,

Mangum B, Mapa P, Martin K, Martin R, Martinez B,

McWhiney S, McLeod M, P, McNeill T, Z, Meenen E,

Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S,

Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L,

Nankervis C, Neal D, Newton N, Nguyen N, Norris S,

Nwaokemelehen O, Okwomou G, Olarinmason A, Pal S, Parks K,

Pasternak S, Paul H, Perez A, Perez L, Pfannkuch C,

Plappier F, Polindexter A, Popovic D, Primus E, Pu L, L,

Piazo M, Quiroz J, Rachlin E, Reeves K, Regier M, A, Reigh R,

Rilly B, Reilly M, Ren Y, Reuter M, Richard S, Riggs F,

Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, J,

Sanders M, Saverly G, Scherer S, Scott G, Shatsman S, Shen H,

Shetty J, Shvartsbeyn A, Sisson I, Sitter C, D, Sma's D,

Sneed A, Sodergren E, Song X, Z, Sorelle R, Soosa J,

Steinle M, Strong R, Sutton A, Svatek A, Tabor P, Taylor C,

Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Umanzi K,

Valae R, Vera V, Villanea D, Waldron J, Walker B, Wang J,

Wang O, Wang S, Warren J, Warren R, Wei K, White F,

Williams G, Willson R, Wleczyk R, Wooden H, Worley K,

Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V,

Yu P, Zhang J, Zhou J, Zhou X, Zhao S, Dunn D, von

Niederhausen A, Weis R, Smith D, R, Holt R, A, Smith H, O,

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

Weinstock G, and Gibbs R, A.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 184532)
AUTHORS Dauphin,S. and Hakenson,W.
TITLE The sequence of Mus musculus BAC clone RP24-459N19
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 184532)
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 184532)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (14-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 184532)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (19-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 184532)
AUTHORS Wilson,R.K.
TITLE Direct Submision
JOURNAL Submitted (02-OCT-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 184532)
AUTHORS Wilson,R.
TITLE Direct Submision
JOURNAL Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 2, 2003 this sequence version replaced gi:27803375.
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0459N19

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu
SOURCE INFORMATION:
The RPCI-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org
NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC021667.
Location/Qualifiers
1. .184532
/organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"
/clone="RP24-459N19"
/clone_id="RPCI-24"
2657. .2749
/rpt_family="L1"
repeat_region
9553. .9657
/rpt_family="Alu"
repeat_region
13735. .13788
/rpt_family="ERV1"
repeat_region
16979. .17050
/rpt_family="B4"
misc_feature
19402. .20528
/note="CpG island (%GC=61.5, o/e=0.81, #CpGs=82)"
27391. .27618
/note="Sequence derived from one plasmid subclone."
35924. .36046
/rpt_family="ERVK"
repeat_region
36203. .36316
/note="Sequence derived from PCR product of project DNA."
44255. .44412
/rpt_family="ERVK"
misc_feature
67067. .68227
/note="CpG island (%GC=61.0, o/e=0.72, #CpGs=74)"
79424. .80411
/note="CpG island (%GC=63.3, o/e=0.94, #CpGs=81)"
84246. .84488
/rpt_family="ERVK"
repeat_region
85290. .85359
/rpt_family="ERV1"
repeat_region
103839. .103921
/rpt_family="ERV1"
repeat_region
107737. .107995
/rpt_family="MER1_type"
112536. .113536
/note="CpG island (%GC=71.3, o/e=0.75, #CpGs=108)"
116901. .116956
/rpt_family="ERV1"
repeat_region
122809. .122934
/rpt_family="MER2_type"
repeat_region
122940. .123121
/rpt_family="L1"
repeat_region
125910. .125991
/rpt_family="MALR"
repeat_region
126174. .126343
/rpt_family="B4"
repeat_region
126376. .126499
/rpt_family="Alu"
repeat_region
126377. .126524
/rpt_family="B4"
repeat_region
127557. .127594
/rpt_family="B2"
repeat_region
128080. .128211
/rpt_family="Alu"
repeat_region
128101. .128291
/rpt_family="B4"
repeat_region
128619. .128690
/rpt_family="B2"
repeat_region
130435. .130504
/rpt_family="ID"
repeat_region
130519. .130771
/rpt_family="Alu"
repeat_region
130530. .130559
/rpt_family="B4"
repeat_region
130560. .130663
/rpt_family="Alu"
repeat_region
130664. .130810
/rpt_family="B4"
repeat_region
130744. .130804
/rpt_family="MALR"
repeat_region
130805. .130890

```
repeat_region /rpt_family="Alu"
130831. .130908 /rpt_family="MALR"
repeat_region 131637. .131710
/rpt_family="I1"
repeat_region 132036. .132173
/rpt_family="MIR"
repeat_region 132229. .132420
/rpt_family="I1"
repeat_region 132589. .132719
/rpt_family="Alu"
repeat_region 132625. .132739
/rpt_family="MALR"
repeat_region 132982. .133110
/rpt_family="Tc2"
repeat_region 133154. .133332
/rpt_family="Alu"
repeat_region 133399. .133593
/rpt_family="B2"
repeat_region 133975. .134204
/rpt_family="B4"
repeat_region 134214. .134325
/rpt_family="Alu"
repeat_region 134638. .134710
/rpt_family="tRNA-Ala-GCA"
repeat_region 134740. .134925
/rpt_family="I1"
repeat_region 135953. .136036
/rpt_family="B4"
repeat_region 136118. .136259
/rpt_family="Alu"
repeat_region 136392. .136569
/rpt_family="I1"
repeat_region 136705. .136859
/rpt_family="ERV.L"
repeat_region 137058. .137195
/rpt_family="Alu"
repeat_region 137500. .137647
/rpt_family="Alu"
repeat_region 137648. .137764
/rpt_family="Alu"
repeat_region 138021. .138147
/rpt_family="MIR"
repeat_region 138278. .138348
/rpt_family="ID"
repeat_region 139137. .139318
/rpt_family="B2"
```

```
Query Match 6.9%; Score 38.8; DB 9; Length 184532;
Best Local Similarity 51.8%; Pred.No. 59;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 364 AGGGAGGCTTGGGAAACCATCTGAAGGCGATGCTTGATTGAGAGGCTGGGCG 423
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 81285 AAGTATGCTAAGGGAACATCTTGAAGGCTTTTTTTTTTTTATGATTTGGGAGCT 81344
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 424 TGGGCTGGGCAAGGCAACAGGCTGAGTCAGAGCCAGAGCAAGGAGCTGCCCAAC 483
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 81345 TAGGTGGGTAGAGGTGTCTATCTTAATACCAATCAAAAGTAGGCAAGGCTAGGCGCTGA 81404
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 484 ACTGCCCGCGCGCTCTGCGATCAGTCTCTCGGCCACCTGAGAACAGCC 533
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 81405 AGTCCCAAGCCTTTCACTCTATTTCTTCTTCAATCTCAGTCCAACC 81454
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 25, 2006, 17:51:42
Job time : 3016.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 15:57:57 ; Search time 472 Seconds
(without alignment)
7963.743 Million cell updates/sec

Title: SEQ1-69C
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtgcgctcttcggaattc 564

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*
- 14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	563	99.8	564	13 ADT97901	Adt97901 Human fib
2	553	98.0	29340	13 ABD33482	Abd33482 Human can
3	553	98.0	57501	14 ADZ13482	Adz13482 Human can
4	38.2	6.8	2161	8 ABX34807	Abx34807 Human md
5	38.2	6.8	2777	11 ADM01893	Adm01893 Human cdn
6	37.2	6.6	712	6 ABK34393	Abk34393 Human cdn
7	37.2	6.6	1957	4 ABK42424	Abk42424 Genomic s
8	37.2	6.6	1957	4 AAL07453	Aal07453 Human rep
9	37.2	6.6	1957	4 ADB60580	Adb60580 Connectiv
10	37	6.6	1785	6 AAS19511	Aas19511 Reference
11	36.6	6.5	550	9 ACH25075	Ach25075 Human adu
12	36.4	6.5	76180	13 ABD33385	Abd33385 Human can
13	36.2	6.4	138203	14 ADZ51725	Adz51725 FR-008 po
14	36	6.4	546	13 ADO56739	Ado56739 Novel can
15	35.8	6.3	704	2 AAO94044	Aag94044 Human ALD
16	35.8	6.3	1263	6 ABO90455	Abg90455 M. capsul
17	35.6	6.3	568	5 ABV51994	Abv51994 Human pro
18	35.2	6.2	954	14 ACL70503	Actl70503 M. xanthu
19	35.2	6.2	4050	14 ACL64141	Actl64141 M. xanthu

20	35	6.2	1338	6 ABL58452	Ab158452 Human pel
21	35	6.2	1338	12 ADP48674	Adp48674 Human pel
22	35	6.2	1797	4 ABK42425	Abk42425 Genomic s
23	35	6.2	1797	4 AAL07455	Aal07455 Human rep
24	35	6.2	1797	9 ADB60581	Adb60581 Connectiv
25	35	6.2	2000	8 ADA71938	Ada71938 Rice gene
26	34.8	6.2	792	13 ADR62531	Adr62531 Cotton cd
27	34.6	6.1	40742	4 AAK68089	Aak68089 Human imm
28	34.6	6.1	215980	6 AAK79886	Aak79886 Human imm
29	34.6	6.1	215980	6 AAL38337	Aal38337 Complemen
30	34.6	6.1	220895	6 AAK84798	Aak84798 Human cdn
31	34.6	6.1	220895	13 ADR52737	Adr52737 Drug ther
32	34.4	6.1	1988	13 ADS56480	Ads56480 Bacterial
33	34.4	6.1	2000	8 ADA71938	Ada71938 Rice gene
34	34.4	6.1	5051	13 ADO39151	Ado39151 Human SNP
35	34.4	6.1	5232	13 ADO39152	Ado39152 Human SNP
36	34.4	6.1	92638	6 ABO88096	Abg88096 Human ost
37	34.2	6.1	756	2 AAX85016	Aax85016 Human sec
38	34.2	6.1	756	6 ACD18942	Actl18942 Novel hum
39	34.2	6.1	756	12 ADG78333	Adg78333 Human sec
40	34.2	6.1	756	12 ADN60624	Adn60624 Human sec
41	34.2	6.1	349980	6 ABO81844	Abg81844 Bifidobac
42	34	6.0	536	10 ADB58842	Adb58842 Minorily
43	34	6.0	593	12 ACH71544	Ach71544 Human gen
44	34	6.0	1028	10 ADD84538	Add84538 121P1 v
45	34	6.0	1028	10 ADD84540	Add84540 121P1 v

ALIGNMENTS

RESULT 1	
ADT97901	
ID	ADT97901 standard; DNA; 564 BP.
XX	
AC	ADT97901;
XX	
DT	27-JAN-2005 (first entry)
XX	
DE	Human fibroblast growth factor (FGF)-3 promoter 5' proximal region.
XX	
KW	Human; ds; fibroblast growth factor-3; FGF-3; promoter; SNP;
KW	single nucleotide polymorphism; cancer; oesophageal cancer; 5'UTR;
KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
KW	oesophageal squamous cell carcinoma.
XX	
OS	Homo sapiens.
XX	
XX	Key
XX	Location/Qualifiers
FT	variation
FT	replace(69,Y)
FT	/*tag= a
FT	/standard_name= "Single nucleotide polymorphism"
FT	/note= "This SNP is specifically claimed in claim 6"
XX	
XX	US2004219582-A1.
XX	
XX	04-NOV-2004.
XX	
XX	11-MAR-2004; 2004US-00798652.
XX	
XX	17-MAR-2003; 2003US-0455689P.
XX	
XX	(GUOY/) GUO Y.
XX	
XX	Guo Y;
XX	
XX	WPI; 2004-794435/78.
XX	
XX	Novel isolated nucleic acid molecule having single nucleotide
XX	polymorphism in upstream untranslated region of fibroblast growth factor-
XX	3 gene; useful for assessing related susceptibility of mammal to cancer.
XX	
XX	Claim 1; SEQ ID NO 1; 15p; English.

Db 3746 CTGGTGCAGGCTCTTGGAAAGTCTGGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTG 3805

Qy 241 GCCGSCATGCCCCCACAAGTTGCACATCACTGCTTTGCAGGGTTGGGCTTCAAGTC 300

Db 3806 GCCGSCATGCCCCCACAAGTTGCACATCACTGCTTTGCAGGGTTGGGCTTCAAGTC 3855

Qy 301 ACAGGGTCCATCCACGTACCAAGCCCAAGTGGCTGCAGAAAGTCCCTCGCAGTCATGAA 360

Db 3866 ACAGGGTCCATCCACGTACCAAGCCCAAGTGGCTGCAGAAAGTCCCTCGCAGTCATGAA 3925

Qy 361 CCAAAGGAGGCTTGGGAAACCAATCTGAAGGGCATGCTTTGATTTAGTAGAGGGTGG 420

Db 3926 CCAAGGAGGCTTGGGAAACCAATCTGAAGGGCATGCTTTGATTTAGTAGAGGGTGG 3985

Qy 421 GGCTGGGCTTGGGCAAGGCCACCAAGGTTGAAGTCAGAGCCAGAGGCAAGAGCTGGTCCC 480

Db 3986 GGCTGGGCTTGGGCAAGGCCACCAAGGTTGAAGTCAGAGCCAGAGGCAAGAGCTGGTCCC 4045

Qy 481 A-GCACTGCCCGCGGCTCTGTGGATGCAGTCTCTTGGCCACCTGAGAAACGCTGTAGA 539

Db 4046 AGGCACCTGCCCGCGGCTCTGTGGATGCAGTCTCTTGGCCACCTGAGAAACGCTGTAGA 4105

Qy 540 GAGGCAGTGGGCTCTTTGGACTTC 564

Db 4106 GAGGCAGTGGGCTCTTTGGACTTC 4130

RESULT 3	
AD213482	
ID	AD213482 standard; DNA; 57501 BP.
XX	
AC	
XX	AD213482;
DT	16-JUN-2005 (first entry)
XX	
DE	Human cancer-associated genomic DNA #85.
XX	
XX	Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
KW	cytostatic; gene; ds.
OS	
XX	Homo sapiens.
XX	
PN	WO2005031001-A2.
PD	
XX	07-APR-2005.
XX	
PF	23-SEP-2004; 2004WO-US031617.
XX	
PR	23-SEP-2003; 2003US-00669920.
XX	
PA	(CHIR) CHIRON CORP.
XX	
P1	Morris DW, Malandro MS;
DR	WPI; 2005-273395/28.
XX	
PT	Nucleic acid array useful for detecting cancer associated nucleic acid,
XX	comprises two or more nucleic acid probes.
PS	
XX	Disclosure; SEQ ID NO 1002; 198bp; English.
CC	
XX	
CC	The invention relates to a nucleic acid array for detecting a cancer
CC	associated (CA) nucleic acid, comprising two or more nucleic acid probes.
CC	The invention also relates to a peptide array comprising two or more
CC	isolated polypeptides encoded by a CA nucleic acid sequence, a compound
CC	that binds to a polypeptide, an isolated antibody or its fragment which
CC	binds to a polypeptide, which is prepared by immunizing a host animal
CC	with a composition comprising the polypeptide or its antigen binding
CC	fragment and collecting cells from the host expressing antibodies against
CC	the antigen or its antigen binding fragment, a composition comprising the
CC	antibody and a carrier, a method of screening for anticancer activity, a
CC	method of detecting a CA nucleic acid, a method of diagnosing cancer, a

method of treating cancer and a method of inhibiting expression of a CA nucleic acid in a cell. The CA nucleic acids are useful for detecting CA nucleic acids. The antibody is useful for detecting the presence or absence of cancer cells in an individual which involves contacting cells from the individual with the antibody and detecting a complex of a CA protein from the cancer cells and the antibody, where the detection of the complex correlates with the presence of cancer cells in the individual. The composition is useful for inhibiting growth of cancer cells in an individual or for delivering a therapeutic agent to cancer cells in an individual. The invention is also useful for diagnosing cancer, for treating cancer and for inhibiting expression of a CA gene in a cell. This sequence represents human cancer-associated genomic DNA of the invention.

Query Match	98.0%	Score 553;	DB 14;	Length 57501;
Best Local Similarity	99.8%	Pred. No. 1.7e-138;		
Matches 564; Conservative	0;	Mismatches	1;	Gaps 1;

Qy	1	GCAGCCCTGCCTCAGAAAACAGAGAGACGAGACA	CACTCAAGGAGACTCACCCCATATG	60
Db	3716	GCACCCCTGCCTCAGAAAACAGAGAGACGAGACA	CACTCAAGGAGACTCACCCCATATG	3775
Qy	61	GCTGAGAGCGAGGAGAGCTCTTGAGGACAGGAC	CGCCGTCAAGTGGGTGACGGCA	120
Db	3776	GCTGAGAGCGAGGAGAGCTCTTGAGGACAGGAC	CGCCGTCAAGTGGGTGACGGCA	3835
Qy	121	GGGGTCTTGGCCATGGTGGGGACA	CAGGGGCTGCATACAGCTTACTCAGTGACAAATCGATCC	180
Db	3836	GGGGTCTTGGCCATGGTGGGGACA	CAGGGGCTGCATACAGCTTACTCAGTGACAAATCGATCC	3895
Qy	181	CTGGTGGCAAGCTCTTGGAAGTCTGGGAAGTGA	AGCAATGTTTCCCTTAAAGAAATGTCGTG	240
Db	3896	CTGGTGGCAAGCTCTTGGAAGTCTGGGAAGTGA	AGCAATGTTTCCCTTAAAGAAATGTCGTG	3955
Qy	241	GCCGGCATGCGCCCCCAACGTTGACAC	TACTACCTCTTGACAGGGTGGGGCTTCCAGTC	300
Db	3956	GCCGGCATGCGCCCCCAACGTTGACAC	TACTACCTCTTGACAGGGTGGGGCTTCCAGTC	4015
Qy	301	ACAGGGTCCCATCCACGTACCAAGCCCAAGTGG	CTGCAGAAAGTCCCTTCGACGTATGAAA	360
Db	4016	ACAGGGTCCCATCCACGTACCAAGCCCAAGTGG	CTGCAGAAAGTCCCTTCGACGTATGAAA	4075
Qy	361	CCAAAGGAGAGCTTGGGAAACCA	CACTCTGGAAGGGCATGGCTTTGATTTAAGTGAAGGGGTGG	420
Db	4076	CCAAAGGAGAGCTTGGGAAACCA	CACTCTGGAAGGGCATGGCTTTGATTTAAGTGAAGGGGTGG	4135
Qy	421	GGCTGGGCTGGGACCAAGGSCAC	AGAGTCTGAATGCAGACCAAGGSCAGAAAGCTGATCCCC	480
Db	4136	GGCTGGGCTGGGACCAAGGSCAC	AGAGTCTGAATGCAGACCAAGGSCAGAAAGCTGATCCCC	4195
Qy	481	A-GCACTGCGCGCGCTTGCATGCA	GTCTTCTGGCCACTTGAAACAGCCTGTGAGA	539
Db	4196	AGGCACTGCGCGCGCTTGCATGCA	GTCTTCTGGCCACTTGAAACAGCCTGTGAGA	4255
Qy	540	GAGGCAAGTGGCGCTTTTTCGAACTTC	564	
Db	4256	GAGGCAAGTGGCGCTTTTTCGAACTTC	4280	

XX	RESULT 4
ABX34807/C	ABX34807 standard; cDNA; 2161 BP.
ID	ABX34807 standard; cDNA; 2161 BP.
XX	
AC	ABX34807;
XX	
DT	13-FEB-2003 (first entry)
XX	
DE	Human mddt cDNA SEQ ID 368.
XX	
MDDT; human; disease detection and treatment molecule polypeptide; KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;	

QY 199 AGCTGGAAGTGAAGATGTTCCATTAGGAAGTGTGCGCCGCGCATGCCCCCA 258
DB 225 AGTGAAGCAGGACCTGTTCTCTCTGATGACACCCAGGAGCCACGAAGCCGATGAC 166
QY 259 CGTTGACACATCAGCTGTTGACAGGTTGG 289
DB 165 CATGACAAAGGCGCCGCTGATGATGAGCAGG 135

RESULT 6
ABK4393/C
ID ABK4393 standard; cDNA, 712 BP.
XX
AC ABK4393;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human cDNA for novel secreted protein, SEQ ID 162.
XX
KW Human; ss; gene; secreted protein; immune deficiency; viral infection;
KW bacterial infection; fungal infection; autoimmune disorder; burn;
KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
KW lymphoid cell deficiency.
XX
OS Homo sapiens.
XX
FN WO200177290-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010295.
XX
PR 06-APR-2000; 2000US-0194941P.
XX
PA (GEMV) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fecthel K, Agoetino MJ, Howes SH, Resnick RJ,
PI Gulukota K, Graham JR;
XX
DR WPI; 2002-179323/23.
XX
XX Six hundred and twenty five polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for treating
PT immune deficiencies and disorders such as autoimmune disorders.
XX
XX Claim 1; Page 114; 339pp; English.

CC The invention relates to 625 polynucleotides which have been derived from
CC a variety of human tissue sources and which encode novel secreted
CC proteins, their complements and sequences that hybridise to them. Also
CC included are a vector comprising the polynucleotide, a host cell
CC transformed with the vector, the proteins encoded by the polynucleotides,
CC antibodies that bind to the proteins and identification of modulators of
CC the proteins or the expression of the polynucleotide. The polynucleotides
CC can be used as probes for the identification and isolation of full length
CC cDNA and genomic DNA. The polynucleotides and proteins can also be used
CC as nutritional supplements. The protein is useful in the treatment of
CC various immune deficiencies and disorders such as viral infections,
CC bacterial infections, fungal infections, autoimmune disorders (e.g.
CC rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and
CC diabetes) and allergic reactions and conditions (e.g. asthma). They are
CC also useful for treating neurodegenerative diseases (e.g. Alzheimer's
CC disease, Parkinson's disease), liver fibrosis, coagulation disorders
CC (e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and
CC tumours. They are also useful for tissue regeneration, for wound healing
CC and in the treatment of burns, incisions and ulcers. The proteins are
CC also useful for regulating haematopoiesis, for treating myeloid or
CC lymphoid cell deficiencies. The present sequence is one of the 625 cDNA
CC sequences encoding a secreted protein

XX SQ Sequence 712 BP, 141 A; 159 C; 205 G; 207 T; 0 U; 0 Other;
Query Match 6.6%; Score 37.2; DB 6; Length 712;
Best Local Similarity 56.6%; Pred. No. 7.1;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCCAAGTTGACACTGCTTGTGACAGGTTGGGGCTTCCAGTCAAGAGGTCCC 310
DB 560 CCGCCCATGCTGGCTGACTAGTTCTCCCTCGAGGGAAGGGGCCGAGGGCTTCCCTCCC 501
QY 311 ATCCAGCTACCAAGCCCAAGTGGCTGCAAGAGTCCCTGCAATCAAAACCAAGGAGG 370
DB 500 AGGGAATCTCCACCAAGAGTGTGACAGCGGGCTGCGCTGAGTCAATCCAGGAACA 441
QY 371 CT 372
DB 440 CT 439

RESULT 7
ABK4242
ID ABK4242 standard; DNA, 1957 BP.
XX
AC ABK4242;
XX
DT 21-MAY-2002 (first entry)
XX
DE Genomic sequence #323 encoding novel human connective tissue polypeptide.
XX
KW Human; connective tissue related disorder; cancer; gene therapy;
KW cytoskeletal; gene; ds.
XX
OS Homo sapiens.
XX
FN WO200155343-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001322.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218280P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226661P.

RESULT 8
AAL07453
ID AAL07453 standard; DNA; 1957 BP.
XX AC AAL07453;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 10141.
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KM cancer; gene therapy; da.
XX OS Homo sapiens.
XX PN WO200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001339.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246613P.
PR 08-NOV-2000; 2000US-0246617P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR	21-SEP-2000	2000US-02342233
PR	21-SEP-2000	2000US-0234274P
PR	25-SEP-2000	2000US-0234937P
PR	25-SEP-2000	2000US-0234938P
PR	26-SEP-2000	2000US-0235484P
PR	27-SEP-2000	2000US-0235834P
PR	27-SEP-2000	2000US-0235836P
PR	29-SEP-2000	2000US-0236337P
PR	29-SEP-2000	2000US-0236337P
PR	29-SEP-2000	2000US-0236338P
PR	29-SEP-2000	2000US-0236339P
PR	29-SEP-2000	2000US-0236802P
PR	02-OCT-2000	2000US-0237037P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0239935P
PR	13-OCT-2000	2000US-0239937P
PR	13-OCT-2000	2000US-0240963P
PR	20-OCT-2000	2000US-0241221P
PR	20-OCT-2000	2000US-0241785P
PR	20-OCT-2000	2000US-0241786P
PR	20-OCT-2000	2000US-0241787P
PR	20-OCT-2000	2000US-0241808P
PR	20-OCT-2000	2000US-0241809P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246534P
PR	08-NOV-2000	2000US-0246535P
PR	08-NOV-2000	2000US-0246536P
PR	08-NOV-2000	2000US-0246577P
PR	08-NOV-2000	2000US-0246582P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	08-NOV-2000	2000US-0246613P
PR	17-NOV-2000	2000US-0249207P
PR	17-NOV-2000	2000US-0249308P
PR	17-NOV-2000	2000US-0249209P
PR	17-NOV-2000	2000US-0249321P
PR	17-NOV-2000	2000US-0249311P
PR	17-NOV-2000	2000US-0249312P
PR	17-NOV-2000	2000US-0249313P
PR	17-NOV-2000	2000US-0249314P
PR	17-NOV-2000	2000US-0249315P
PR	17-NOV-2000	2000US-0249365P
PR	17-NOV-2000	2000US-0249367P
PR	17-NOV-2000	2000US-0249371P
PR	17-NOV-2000	2000US-0249372P
PR	17-NOV-2000	2000US-0249373P
PR	17-NOV-2000	2000US-0249374P
PR	17-NOV-2000	2000US-0249375P
PR	17-NOV-2000	2000US-0250160P
PR	01-DEC-2000	2000US-0250391P
PR	05-DEC-2000	2000US-0251030P
PR	05-DEC-2000	2000US-0251988P
PR	05-DEC-2000	2000US-0256719P
PR	06-DEC-2000	2000US-0251579P
PR	06-DEC-2000	2000US-0251586P
PR	08-DEC-2000	2000US-0251688P
PR	08-DEC-2000	2000US-0251697P
PR	08-DEC-2000	2000US-0251899P

PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0255678P.
PR	17-JAN-2001;	2001US-00764847.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Rosen CA, Ruben SM, Barash SC;	
PI		
DR	WPI; 2003-634869/60.	
XX	P-PSDB; ADB59454.	
PT	New connective tissue-related polypeptides and polynucleotides, useful	
PT	for treating, preventing and/or prognosing e.g. disorders of connective	
PT	tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or	
PT	neoplasias.	
XX		
PS	Disclosure; SEQ ID NO 1311; 248bp; English.	
XX		
CC	The invention describes an isolated nucleic acid molecule (1), which	
CC	comprises a sequence that is at least 95 % identical to a connective	
CC	tissue-related polynucleotide encoding connective tissue antigens (CTA).	
CC	The polypeptide or polynucleotide is useful for preventing, treating, or	
CC	ameliorating medical conditions in a mammal. The connective tissue	
CC	polypeptides, polynucleotides and antibodies are particularly useful for	
CC	treating, preventing and/or prognosing disorders of connective tissues	
CC	(e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,	
CC	scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or	
CC	neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, or Parkinson's disease), cardiovascular diseases	
CC	(e.g. atherosclerosis, myocarditis or cardiopulmonary bypass	
CC	complications), autoimmune diseases (e.g. systemic lupus erythematosus,	
CC	rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.	
Query Match	6.6%; Score 37.2; DB 9; Length 1957;	
Best Local Similarity	56.6%; Pred. No. 9.3;	
Matches	69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;	
OY	251 CCCCCACGTTGACACTCACTGCTTGGAGGGTGGGGCTTCCAGTCACAGGGTCCC 310	
Db	165 CCGCCCATGCTGGCTCACTTCTCTGCGCTGGAGGGAAGGGCCGAGGGCTTCTGTCCC 224	
OY	311 ATCCACGTACAGCCCAAGTGGCTGTCAGAAAGTCCCTCGCAGTCATGAAACCAAGGAGAG 370	
Db	225 AGGGAACCTCCACACACAGAGACTGTGCAGACGGGCTGGCAGTCACTGCTCCAGAAC 284	
OY	371 CT 372	
Db	285 CT 286	
RESULT 10		
AS1511		
XX	AS1511 standard; DNA; 1785 BP.	
XX		
AC	AS1511;	
XX		
DT	26-MAR-2002 (first entry)	
XX		
XX	Reference sequence for human MPL gene exons 9-10.	
DE		
XX		
XX	Human; single nucleotide polymorphism; SNP; MPL; chromosome 1p34;	
KW	myeloproliferative leukaemia virus oncogene; haplotyping; genotyping;	
KW	congenital amegakaryocytic thrombocytopoenia; CAMT; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	exon	614..773
FT		/*tag= a
FT		/number= 9
FT	intron	774..1027
FT		/*tag= b

PN WO2004058146-A2.
XX
PD 15-JUL-2004.
XX
XX 15-DEC-2003; 2003WO-US040081.
XX
XX 17-DEC-2002; 2002US-00322281.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Melandro MS;
XX WPI; 2004-499109/47.
XX
XX Novel human cancer associated protein encoded within open reading frame
XX of cancer associated gene, useful as targets for diagnosing cancer.
XX
XX Claim 16; SEQ ID NO 492; 182pp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
XX associated (CA) nucleic acids encoding them. The invention also relates
XX to a method for treating cancers involving administering to a patient an
XX inhibitor of CAP, and a method of screening for anticancer activity in a
XX potential drug involving providing a cell that expresses a CA gene,
XX contacting a tissue sample derived from a cancer cell with an anticancer
XX drug candidate and monitoring the effect of the anticancer drug candidate
XX on expression of the CA gene. The CAP proteins are useful for detecting
XX cancer associated with expression of a CAP protein in a test cell sample
XX and for screening for a bioactive agent capable of modulating the
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX cancer, involving determining the expression of a CA nucleic acid in a
XX tissue. This sequence represents a human CA gene of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 76180 BP; 17603 A; 18463 C; 19657 G; 20457 T; 0 U; 0 Other;
XX
XX Query Match 6.5%; Score 36.4; DB 13; Length 76180;
XX Best Local Similarity 59.8%; Pred. No. 41;
XX Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
XX
QY 25 GGACGCAGACACTCAGCGTACTCACCCTGCTGAGAGCGAGCGCTCTGA 84
XX
Db 75766 GGGGGCAGCCCTGGCATGTACTTTGGCAGATGAGAGAAACACAGAGCAGCGAGA 75707
XX
QY 85 GGCAGGGCCAGGGCAGCCCTCAGGTGGGTGACGGCAGGGGTC 126
XX
Db 75706 GGAAGGGAGAGCGAGCCAGCAGCAGAGGTGACAGGTGACATC 75665
XX
RESULT 13
AD251725
ID AD251725 standard; DNA; 138203 BP.
XX
AC AD251725;
XX
XX 14-JUL-2005 (first entry)
XX
XX FR-008 polypeptide gene cluster, SEQ ID 1.
XX
XX Antifungal; antiparasitic; polypeptide; gene; db.
XX
XX Streptomyces sp.; FR-008.
XX
XX Key Location/Qualifiers
XX FT complement(574..1950)
XX CDS
XX /product= "FAD-dependent monooxygenase, FSCO, SEQ ID 20"
XX /tag= a
XX /product= "ADC lyase, PabC, SEQ ID 22"
XX /tag= b
XX complement(3150..3818)
XX CDS

FT /tag= c
FT /product= "Regulator protein, FscR1, SEQ ID 10"
FT complement(4377..7205)
FT CDS
FT /tag= d
FT /product= "Regulator protein, FscR11, SEQ ID 11"
FT complement(7210..10320)
FT CDS
FT /tag= e
FT /product= "Regulator protein, FscR11f, SEQ ID 12"
FT complement(10298..13315)
FT CDS
FT /tag= f
FT /product= "Regulator protein, FscR1V, SEQ ID 13"
FT 13522..14898
FT CDS
FT /tag= g
FT /product= "Glycosyltransferase, FscW1, SEQ ID 17"
FT 14953..16011
FT CDS
FT /tag= h
FT /product= "GDP-ketosugar aminotransferase, FscM11, SEQ ID
FT 18"
FT 16106..17287
FT CDS
FT /tag= i
FT /product= "Cytochrome P-450 monooxygenase, FscP, SEQ ID
FT 14"
FT 17334..17528
FT CDS
FT /tag= j
FT /product= "Ferredoxin protein, FscFE, SEQ ID 15"
FT 17556..18413
FT CDS
FT /tag= k
FT /product= "Thioesterase, FscTE, SEQ ID 16"
FT 18610..20781
FT CDS
FT /tag= l
FT /product= "ADC synthase, PabAB, SEQ ID 21"
FT 20927..26158
FT CDS
FT /tag= m
FT /product= "PKS FscA, SEQ ID 2"
FT 26333..27340
FT CDS
FT /tag= n
FT /product= "ABC transporter, FscTI, SEQ ID 8"
FT 27561..28280
FT CDS
FT /tag= o
FT /product= "ABC transporter, FscT11, SEQ ID 9"
FT complement(28983..60860)
FT CDS
FT /tag= p
FT /product= "PKS FscC, SEQ ID 3"
FT complement(60962..77587)
FT CDS
FT /tag= q
FT /product= "PKS FscB, SEQ ID 4"
FT /product= "PKS FscF, SEQ ID 5"
FT /tag= r
FT /product= "PKS FscF, SEQ ID 5"
FT complement(84170..107485)
FT CDS
FT /tag= s
FT /product= "PKS FscE, SEQ ID 6"
FT complement(107496..136148)
FT CDS
FT /tag= t
FT /product= "PKS FscD, SEQ ID 7"
FT complement(136558..137766)
FT CDS
FT /tag= u
FT /product= "GDP-mannose-4,6-dehydratase, FscM111, SEQ ID
FT 19"
XX
XX US2005089884-A1.
XX
XX 28-APR-2005.
XX
XX 06-APR-2004; 2004US-00819386.
XX
XX 23-OCT-2003; 2003KR-00074035.
XX
XX (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX (UTSH-) UNITV SHANGHAI JIAOTONG.
XX
XX Lee SY, Deng Z, Chen S, Jeong KJ, Zhou X;
XX PI

[illegible]

Db 550 GGATCTTCAGGGGGCCCTCCACGACGACACCGACCGGC 512

Search completed: March 25, 2006, 16:28:07
Job time : 476 secs

This Page Blank (uspto)

ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 22)
AUTHORS Shen,S., Schedl,A. and Harmar,A.J.
TITLE Vector
JOURNAL Patent: JP 2001525168-A 21 11-DEC-2001;
MEDICAL RESEARCH COUNCIL
OS Artificial Sequence
PN JP 2001525168-A/21
PD 11-DEC-2001
PF 27-NOV-1998 JP 2000523326
PR 28-NOV-1997 GB 9725311.6,28-NOV-1997 GB 9725313.2 PR
20-MAR-1998 GB 9806072.6,05-NOV-1998 GB 9824275.3 PI
SANBING SHEN,ANDREAS SCHEDL,ANTHONY JOHN HARMAR PC
C12N15/09, C12N15/00
CC Description of Artificial Sequence: Primer
FH Key Location/Qualifiers
FT source 1..22
FT Location/Qualifiers
1..22 /organism='Artificial Sequence'.
FEATURES
source 1..22
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 69.1%; Score 15.2; DB 6; Length 22;
Best Local Similarity 85.0%; Pred. No. 4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCTCAGCTCCAGCCACAT 21
1 TTCATCAGCTCCATCCACAT 20

Db 1 TTCATCAGCTCCATCCACAT 20

RESULT 3
AX003404 22 bp DNA linear PAT 24-AUG-2000
LOCUS Sequence 21 from Patent WO928449.
DEFINITION AX003404
ACCESSION AX003404
VERSION AX003404.1 GI:9927208
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Schedl,A. and Harmar,A.J.
TITLE Vectors
JOURNAL Patent: WO 928449-A 21 10-JUN-1999;
SCHIDL ANDREAS (DE); HARMAR ANTHONY JOHN (GB)
FEATURES
source 1..22
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"

ORIGIN
Query Match 69.1%; Score 15.2; DB 6; Length 22;
Best Local Similarity 85.0%; Pred. No. 4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCTCAGCTCCAGCCACAT 21
1 TTCATCAGCTCCATCCACAT 20

Db 1 TTCATCAGCTCCATCCACAT 20

RESULT 4
A57518 24 bp DNA linear PAT 03-MAR-1998
LOCUS Sequence 10 from Patent WO9632483.
DEFINITION A57518
ACCESSION A57518
VERSION A57518.1 GI:3713376

KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1
AUTHORS Masucci,M.G.
TITLE IMMUNE-EVADING PROTEINS
JOURNAL Patent: WO 9632483-A 10 17-OCT-1996;
MASUCCI MARIA GRAZIA (SE)
COMMENT Other publication AU 5284296 961030.
FEATURES
source 1..24
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 66.4%; Score 14.6; DB 6; Length 24;
Best Local Similarity 81.0%; Pred. No. 7.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
4 CACCCGACCTCCAGCTCCAT 24

Db 4 CACCCGACCTCCAGCTCCAT 24

RESULT 5
AR052984 24 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 16 from patent US 5833991.
DEFINITION AR052984
ACCESSION AR052984
VERSION AR052984.1 GI:5977846
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Masucci,M.G.
TITLE Glycine-containing sequences conferring inviability to the immune system
JOURNAL Patent: US 5833991-A 16 10-NOV-1998;
FEATURES
source 1..24
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 66.4%; Score 14.6; DB 6; Length 24;
Best Local Similarity 81.0%; Pred. No. 7.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
4 CACCCGACCTCCAGCTCCAT 24

Db 4 CACCCGACCTCCAGCTCCAT 24

RESULT 6
AR306619 30 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 10 from patent US 6548642.
DEFINITION AR306619
ACCESSION AR306619
VERSION AR306619.1 GI:31696821
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums
JOURNAL Patent: US 6548642-A 10 15-APR-2003;
Ohio University; Athens, OH
FEATURES
source 1..30

RESULT 11				
AR412132	AR412132	30 bp	DNA	linear
LOCUS	Sequence	38	from patent US 6639050.	
DEFINITION	AR412132			
ACCESSION	AR412132			
VERSION	AR412132.1	GI:4016776		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
	Unclassified.			

REFERENCE	1 (bases 1 to 30)
AUTHORS	Kieliszewski, M.J.
TITLE	Synthetic genes for plant gums and other hydroxyproline-rich glycoproteins
JOURNAL	Patent: US 6639050-A 38 28-OCT-2003; Ohio University; Athens, OH
FEATURES	location/Qualifiers
source	1..30 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Query Match	66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity	81.0%; Pred. No. 7.4e+04;
Matches	17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	1 CTCCTCACCTCCAGCCACAT 21
DB	5 CACCTTCACCTCCACCCCAT 25
RESULT 12	
BD174200/c	21 bp DNA linear PAT 18-FEB-2003
LOCUS	BD174200
DEFINITION	Periplastic converting agent.
ACCESSION	BD174200
VERSION	BD174200.1 GI:28415537
KEYWORDS	WO 02066049-A/46.
SOURCE	synthetic construct
ORGANISM	synthetic construct other sequences: artificial sequences.
REFERENCE	1 (bases 1 to 21)
AUTHORS	Hikichi, Y., Shintani, Y. and Matsui, H.
TITLE	Periplastic converting agent
JOURNAL	Patent: WO 02066049-A 46 29-AUG-2002; TAKEDA CHEMICAL INDUSTRIES LTD, YUKIKO HIKICHI, YASUSHI SHINTANI, HIDEKI MATSUI
COMMENT	OS Artificial Sequence PN WO 02066049-A/46 PD 29-AUG-2002 PF 21-FEB-2002 WO 2002JP001536 PR 23-FEB-2001 JP 01P 09450 PI YUKIKO HIKICHI, YASUSHI SHINTANI, HIDEKI MATSUI PC A61K38/17, A61K31/711, A61K48/00, A61P43/00, A61P21/00, A61P21/04, PC A61P15/00, PC C12N15/12, C07K14/47 CC Primer FH Key location/Qualifiers FT source 1..21 FT /organism='Artificial Sequence'.
FEATURES	Location/Qualifiers
source	1..21 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
ORIGIN	
Query Match	65.5%; Score 14.4; DB 6; Length 21;
Best Local Similarity	93.8%; Pred. No. 9.4e+04;
Matches	15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 CTCCTCACCTCCAGC 16
DB	18 CTCCTCACCTCCATC 3
RESULT 13	
BD185155/c	21 bp DNA linear PAT 17-JUN-2003
LOCUS	BD185155
DEFINITION	Cell differentiating agent.
ACCESSION	BD185155
VERSION	BD185155.1 GI:31877355
KEYWORDS	JP 2002356438-A/46.

ORIGIN	source	synthetic construct
REFERENCE	synthetic construct	other sequences; artificial sequences.
AUTHORS	1 (bases 1 to 21)	Hikichi,Y., Shintani,Y. and Matsui,H.
TITLE	Cell differentiating agent	Patent: JP 200256438-A 46 13-DEC-2002;
JOURNAL	TAKEDA CHEMICAL INDUSTRIES LTD	
COMMENT	OS Artificial Sequence	
	PN JP 200256438-A/46	
	PD 13-DEC-2002	
	PF 21-FEB-2002 JP 2002044741	
	PI YUKIKO HIKICHI,YASUSHI SHINTANI,HIDEKI MATSUI PC	
	A61K38/00,A61K31/7088,A61P15/00,A61P21/04,A61P35/00,C12N15/09// PC	
	C07K4/525	
	PC C1201/68,A61K37/02,C12N15/00	
	CC Primer	
	FM Key	
	FT source	1..21
	FT	/organism='Artificial Sequence'.
FEATURES	Location/Qualifiers	
source	1..21	
	/organism="synthetic construct"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32630"	
	/note="Description of Artificial Sequence: PCR Primer	
	sequence"	
ORIGIN		
Query Match	65.5%; Score 14.4; DB 6; Length 21;	
Best Local Similarity	93.8%; Pred. No. 9.4e+04;	
Matches	15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
CY	1 CTCCTCACCCTCCAGC 16	
Db	18 CTCCTCACCCTCCATC 3	
RESULT 14		
LOCUS	AX817728	22 bp DNA linear PAT 10-DEC-2003
DEFINITION	Sequence 476 from Patent WO02081517.	
ACCESSION	AX817728	
VERSION	AX817728.1 GI:39722920	
KEYWORDS		
SOURCE		
ORGANISM	synthetic construct	
REFERENCE	synthetic construct	
AUTHORS	other sequences; artificial sequences.	
	1	
	Deriastofaro,M.F., Padigaru,M., Miller,C., Tchernen,V., Zhong,H.,	
	Zhong,M., Anderson,D., Ballinger,R., Gerlach,V., Spytek,K.A.,	
	Rastelli,L., Kehruda,R., Guo,X., Zernusen,B., Andrew,D., Meres,P.,	
	Patturajan,M., Burgess,C.E., Eisen,A., Wolenc,A., Baumgartner,J.,	
	Shimkets,R.A., Gusev,V., Vernet,C.A., Taupier,R.J., Pena,C.,	
	Shenoy,S., Li,L., Casman,S., Boligog,F., Fernandes,E., Smithson,G.,	
	Malvankar,U., Tallion,B. and Liu,X.	
	Novel polypeptides and nucleic acids encoded thereby	
	Patent: WO 02081517-A 476 17-OCT-2002;	
	Curagen Corporation (US)	
	Location/Qualifiers	
	1..22	
	/organism="synthetic construct"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:32630"	
	/note="Description of Artificial Sequence: PCR Primer	
	sequence"	
ORIGIN		
Query Match	65.5%; Score 14.4; DB 6; Length 22;	
Best Local Similarity	93.8%; Pred. No. 9.4e+04;	
Matches	15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
CY	1 CTCCTCACCCTCCAGC 16	
Db	1 CTCCTCACCCTCCATC 16	

RESULT 15
A57517/c

LOCUS A57517 24 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 9 from Patent WO9632483.
ACCESSION A57517
VERSION A57517.1 GI:3713375

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT
FEATURES
source

ORIGIN
Query Match 64.5%; Score 14.2; DB 6; Length 24;
Best Local Similarity 84.2%; Pred. No. 1.2e+05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Search completed: March 25, 2006, 14:43:29
Job time : 962 secs

This Page Blank (uspto)

GenCore version 5.1.7
(c) 1993 - 2006 Biocompare

using bw model

March 25, 2006, 13:22:55 ; Search time 372.5 Seconds
(without alignments)
393.619 Million cell updates

US-10-798-652-6

1 ctccctcaccctccagccaatg 22

IDENTITY_NUC
SECRET 10-A SECRET 1 A

4996997 beqs, 3332346308 residues

white satisfying chosen parameters: 4138570

length: 0

Minimum Match 0%

Listing first 45 summaries

N_GeneSeq_21: *

```

1: _ genebegqn1980s:*
2: genebegqn1990s:*
3: genebegqn2000s:*
4: genebegqn2001as:*
5: genebegqn2001bs:*
6: genebegqn2002as:*
7: genebegqn2002bs:*
8: genebegqn2003as:*
9: genebegqn2003bs:*
10: genebegqn2003cs:*
11: genebegqn2003ds:*
12: genebegqn2004as:*
13: genebegqn2004bs:*
14: genebegqn2005s:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	22	13	ADT97906	Adt97906 Human FGF1
2	20.4	92.7	22	13	ADT97907	Adt97907 Human FGF1
C 3	15.6	70.9	25	10	ADB67764	ADB67764 HIV-1 gp120
4	15.2	69.1	22	12	AAK57861	AAK57861 PCR primer
5	15.2	69.1	22	5	AAK65242	AAK65242 Reverse P
6	15.2	69.1	25	9	ACH62883	ACH62883 DNA target
7	15.2	69.1	25	9	ACH63009	ACH63009 DNA target
8	14.6	66.4	24	2	AAV55821	AAV55821 Multimer1
C 9	14.6	66.4	29	14	AAV50944	AAV50944 PCR primer
10	14.6	66.4	30	2	AAK27663	AAK27663 DNA encod
11	14.6	66.4	30	6	ABE51730	ABE51730 HRP re la
12	14.6	66.4	30	6	ABE51739	ABE51739 Hydroxypr
13	14.6	66.4	30	13	ADU09143	ADU09143 Ser-Hyp4
14	14.6	66.4	30	13	ADU09116	ADU09116 Hydroxypr
C 15	14.4	65.5	21	6	AAI49623	AAI49623 Tumour di
C 16	14.4	65.5	21	12	ADM69088	ADM69088 Human TAT
17	14.4	65.5	21	13	ADST74003	ADST74003 A. gossyp
18	14.4	65.5	22	8	ABT33560	ABT33560 NOV forwa
19	14.2	64.5	24	2	AAQ40923	AAQ40923 Probe Pol

C	20	14.2	64.5	24	2	AA739970	Aac59970	Minimal m
C	21	14.2	64.5	24	2	AA558820	Aa558820	Multimeri
C	22	14.2	64.5	28	10	AB224297	Ab224297	Human MI
C	23	14.2	64.5	30	4	AAH21949	Aah21949	Mouse dig
C	24	14.2	64.5	30	12	ADK15423	Adk15423	PCR prim
C	25	14.2	64.5	30	12	AD062731	Ad62731	Human in
C	26	14	63.6	21	13	ADU43700	Ad43700	Knock-do
C	27	14	63.6	24	6	ABK14322	Abk14322	Human P2Y
C	28	14	63.6	25	9	AC149317	Ac149317	Human nic
C	29	13.8	62.7	17	11	AD147245	Ad147245	Human MO
C	30	13.8	62.7	19	13	ADK15645	Adk15645	Human DN
C	31	13.8	62.7	20	3	AA224490	Aa224490	Wheat sol
C	32	13.8	62.7	20	14	AEA51146	Aea51146	pyranalol
C	33	13.8	62.7	21	10	ADG30065	Adg30065	IL2-targ
C	34	13.8	62.7	21	10	ADG30061	Adg30061	IL2-targ
C	35	13.8	62.7	21	10	ADG30057	Adg30057	IL2-targ
C	36	13.8	62.7	21	12	ADG44286	Adg44286	Swine Mx
C	37	13.8	62.7	23	10	ADG29531	Adg29531	IL2-targ
C	38	13.8	62.7	23	10	ADG30053	Adg30053	IL2-targ
C	39	13.8	62.7	24	12	ADP32021	Adp32021	Probe #3
C	40	13.8	62.7	25	2	AAK81755	Aak81755	PCR prime
C	41	13.8	62.7	27	14	AD85535	Ad85535	Interleu
C	42	13.6	61.8	20	4	AA507411	Aa507411	PCR prime
C	43	13.6	61.8	20	4	AA21259	Aa21259	Rat PTPIB
C	44	13.6	61.8	20	6	ABK5234	Abk5234	Rat PTPIB
C	45	13.6	61.8	20	6	ABK37403	Abk37403	Rat PTPIB

ALIGNMENTS

XX	RESULT 1
ADT97906	ID ADT97906 standard; DNA; 22 BP.
AC	ADT97906;
DT	27-JAN-2005 (first entry)
DE	Human FGF-3 promoter 5' proximal region T allele probe.
KM	Human; sb; fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer
KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
KM	oesophageal squamous cell carcinoma; PCR; probe; SNP;
XX	single nucleotide polymorphism.
OS	Homo sapiens.
Key	Location/Qualifiers
FT	modified_base 1
FT	/tag= a
FT	/mod_base= OTHER
FT	/note= "modified with Vic (not defined)"
FT	22
FT	/tag= b
FT	/mod_base= OTHER
FT	/note= "G is covalently linked to a TAMRA moiety (6-carbotetramethyl-rhodamine)"
PN	US2004219582-A1.
PD	04-NOV-2004.
Pf	11-MAR-2004; 2004US-00798652.
PR	17-MAR-2003; 2003US-0455689P.
PA	(GUOY/) GUO Y.
PI	Guo Y;
WPI	2004-794435/78.

PT Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
PS Claim 34; SEQ ID NO 6; 15bp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Taqman PCR probe used to genotype
CC individuals for the T SNP in the UTR of FGF-3.
SQ Sequence 22 BP; 4 A; 12 C; 2 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 22; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTCACCCTCCAGCCACATG 22
Db 1 CTCCTCACCCTCCAGCCACATG 22
RESULT 2
ADT97907
ID ADT97907 standard; DNA; 22 BP.
XX
AC ADT97907;
XX
DT 27-JUN-2005 (first entry)
XX
DE Human FGF-3 promoter 5' proximal region C allele probe.
XX
KW Human; ss: fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer;
KW breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
KW oesophageal squamous cell carcinoma; PCR; probe; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT modified_base 1
FT /*tag= a
FT /mod_base= OTHER
FT /note= "C is covalently linked to a 6-carbo fluorescein
FT moiety (FAM) "
FT 22
FT modified_base
FT /*tag= b
FT /mod_base= OTHER
FT /note= "G is covalently linked to a TAMRA moiety (6-
FT carbocetramethyl-rhodamine) "

XX
PN US2004219582-A1.
XX
XX 04-NOV-2004.
PD
XX
XX 11-MAR-2004; 2004US-00798652.
PF
XX
XX 17-MAR-2003; 2003US-0455689P.
PR
XX
XX (GUOY/) GUO Y.
PA
XX
XX Guo Y;
PI
XX
XX WPI; 2004-794435/78.
DR
XX
PT Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
PS Claim 20; SEQ ID NO 7; 15bp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Taqman PCR probe used to genotype
CC individuals for the C SNP in the UTR of FGF-3.
XX
SQ Sequence 22 BP; 3 A; 12 C; 3 G; 4 T; 0 U; 0 Other;
Query Match 92.7%; Score 20.4; DB 13; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCACCCTCCAGCCACATG 22
Db 1 CTCCTCACCCTCCAGCCACATG 22
RESULT 3
ADB67764/C
ID ADB67764 standard; DNA; 25 BP.
XX
AC ADB67764;
XX
XX 04-DEC-2003 (first entry)
XX
DT HIV-1 gp120 optimised sequencing primer g120c-SF4.
XX
DE HIV-1 gp120 optimised sequencing primer g120c-SF4.
XX
KW HIV; Tac; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV;
KW primer; ss.
XX

OS Synthetic.
 OS Human immunodeficiency virus 1.
 XX
 PN W02003011334-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 26-JUL-2002; 2002WO-EP008343.
 XX
 PR 27-JUL-2001; 2001GB-00018367.
 XX
 PA (GLAX) GLAXOSMITHKLINE BIOLOGICALS SA.
 XX (GLAX) GLAXO GROUP LTD.
 XX
 PI Erc1 PF, Tite JP, Van Wely CA, Voss G;
 XX
 DR WPI; 2003-239474/23.
 XX
 PT Use of an HIV Tat, Nef or gp120 protein or polynucleotide in
 PT manufacturing a vaccine for a prime-boost delivery for the prophylactic
 PT or therapeutic immunization of humans against HIV.
 XX
 PS Example 15; Fig 21; 108pp; English.
 XX
 CC The present invention describes the use of an HIV Tat protein or
 CC polynucleotide, an HIV Nef protein or polynucleotide, or an HIV Tat
 CC protein or polynucleotide linked to an HIV Nef protein or polynucleotide,
 CC and an HIV gp120 protein or polynucleotide in manufacturing a vaccine for
 CC a prime-boost delivery for the prophylactic or therapeutic immunisation
 CC of humans against HIV. The protein or polynucleotide is delivered via a
 CC bombardment approach. Also described: (1) a recombinant DNA molecule
 CC comprising a Nef and/or Tat and/or gp120 gene in a vector in which the
 CC gene of interest is inserted 3' to an enhanced HCMV IE1 promoter; (2)
 CC particles, preferably gold particles, coated with recombinant DNA
 CC comprising a Nef and/or Tat and/or gp120 gene in a vector; and (3) a kit
 CC comprising at least two different vaccine compositions including: (a) a
 CC composition comprising particles coated with DNA encoding gp120 and nef
 CC and/or tat or nef/tat; and (b) a composition comprising gp120 and nef
 CC and/or tat or nef/tat DNA or proteins, where the DNA or proteins are not
 CC coated onto the particles. The HIV proteins or polynucleotides are used
 CC to produce vaccines having anti-HIV activity. The recombinant DNA
 CC molecule can be used in manufacturing an HIV vaccine for the prophylactic
 CC or therapeutic immunisation of humans. The present sequence for the represents an
 CC oligonucleotide primer for sequencing optimised gp120, which is used in
 CC an example from the present invention.
 XX
 SQ Sequence 25 BP; 6 A; 4 C; 12 G; 3 T; 0 U; 0 Other;
 Query Match 70.9%; Score 15.6; DB 10; Length 25;
 Best Local Similarity 81.8%; Pred. No. 9.1e+03;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CTCCTGACCTCGAGCCACATG 22
 DB 22 CTTCCACCTCCCTGCGCACATG 1
 RESULT 4
 AAX57861
 ID AAX57861 standard; DNA; 22 BP.
 XX
 AC AAX57861;
 XX
 DT 15-JUL-1999 (first entry)
 XX
 DE PCR primer used in construction of yeast artificial chromosome.
 XX
 KM YAC; yeast artificial chromosome; PCR primer; sexual dysfunction;
 KM reporter gene; transgenic mammal; therapy; circadian function;
 KM sleep disorder; eating disorder; premenstrual syndrome; birth defect;
 KM autoimmune disorder; ss.
 XX
 OS Synthetic.

XX
 XX GB2331752-A.
 XX
 PN 02-JUN-1999.
 XX
 PD 27-NOV-1998; 98GB-00026126.
 XX
 PF 28-NOV-1997; 97GB-00025311.
 XX
 PR 28-NOV-1997; 97GB-00025313.
 XX
 PR 20-MAR-1998; 98GB-00006072.
 XX
 PR 05-NOV-1998; 98GB-00024275.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Shen S, Schedl A, Harmar AJ;
 XX
 DR WPI; 1999-290603/25.
 XX
 PT New reporter gene labeled YAC vectors and transgenic mammals used for
 PT screening potential active agents.
 XX
 PS Disclosure; Page 56; 98pp; English.
 XX
 CC This sequence represents a PCR primer used in the construction of a yeast
 CC artificial chromosome of the invention. The yeast artificial chromosome
 CC (YAC) vectors contain a reporter gene and transgenic mammals produced
 CC using them may be used to screen for an agent affecting nucleotide
 CC expression and gives easier monitoring of in vivo expression. The vector
 CC is used in the production of transgenic mammals for testing potential
 CC pharmaceutical or veterinary agents. pYMA4 is used to amplify YAC. The
 CC assay may be used to screen for agents useful in treatment of disturbance
 CC of circadian function, sleep disorders, eating disorders, premenstrual
 CC syndrome, autoimmune disorders, birth defects in women and/or sexual
 CC dysfunction. The agents thus detected may be used for treatment of
 CC disorders related to the expression pattern of a nucleotide such as those
 CC above. The vectors have more concentrated YAC DNA, which allows better
 CC and more reliable gene transfer. The presence of a reporter gene allows
 CC easy monitoring of in vivo expression and the vectors allow for gene
 CC overexpression (3-5 fold) and easy site determination. The pYMA4
 CC amplification vector does not contain the thymidine kinase gene, which
 CC causes male infertility in transgenic mice
 XX
 SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
 Query Match 69.1%; Score 15.2; DB 2; Length 22;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 TCCCTGACCTCGAGCCACAT 21
 DB 1 TTGATGACCTCGATCCACAT 20
 RESULT 5
 AAC85242
 ID AAC85242 standard; DNA; 22 BP.
 XX
 AC AAC85242;
 XX
 DT 22-MAR-2001 (first entry)
 XX
 DE Reverse primer D 26376 for determining size of YAC insert.
 XX
 KM Internal ribosomal entry site; IRS; yeast artificial chromosome; YAC;
 KM vector; centromere; telomere; origin of replication; transgenic;
 KM circadian function; sleep disorder; eating disorder;
 KM premenstrual syndrome; autoimmune disease; birth defect;
 KM sexual dysfunction; serotonin transporter; VIP2 receptor; SERT; VIPR2;
 KM polymerase chain reaction; PCR; primer; amplify; YAC 3508/D6;
 KM YAC HSC7526/V12; ss.
 XX
 OS Synthetic.

PN GB2350613-A.
 XX
 PD 06-DEC-2000.
 XX
 PF 17-AUG-2000; 2000GB-00020335.
 XX
 PR 28-NOV-1997; 97GB-00025311.
 PR 28-NOV-1997; 97GB-00025313.
 PR 20-MAR-1998; 98GB-00006072.
 PR 05-NOV-1998; 98GB-00024275.
 PR 27-NOV-1998; 98GB-00026126.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Shen S, Schedl A, Harmar AJ;
 XX
 DR WPI; 2001-034098/05.
 XX
 PT Transgenic organism for identifying potential therapeutic agents able to
 PT modulate gene expression, comprises a yeast artificial chromosome vector.
 XX
 PS Example; Page 54; 93pp; English.

CC The sequences given in AAC85227-50 are primers which were used to
 CC determine the size of the integrated YAC 35D8/D6 and YAC HSC78526/V12
 CC constructs in a transgenic founder animal. The constructs were prepared
 CC from novel yeast artificial chromosome (YAC) vectors each of which
 CC comprises a centromere, two telomeres, at least one origin of
 CC replication, an internal ribosomal entry site (IRES), and a selection
 CC gene that is specifically removable from the vector. The resulting YAC
 CC are used to produce transgenic organisms for use in screening for agents
 CC that can affect the expression pattern of a nucleotide sequence of
 CC interest (NOI) or the activity of its expression product. The identified
 CC agents are potentially useful as pharmaceutical and veterinary agents for
 CC treating disorders of circadian function; sleep or eating disorders;
 CC premenstrual syndrome; autoimmune diseases; birth defects in women and/or
 CC sexual dysfunction, and also as lead compounds for developing agents with
 CC other activities. YAC can also be used for expression, regulation and/or
 CC functional studies on NOI, in combination with other NOI, compounds or
 CC compositions. The new vectors provide high YAC copy numbers and allow
 CC easy monitoring (in vivo) of the expression pattern of NOI;
 CC (over)expression of NOI and a reporter gene; and determination of the
 CC sites where NOI is expressed. Incorporation of IRES allows expression of
 CC at least two nucleic acid sequences (e.g. NOI plus a reporter of more
 CC than one NOI)
 XX
 XX
 SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;

Query Match 69.1%; Score 15.2; DB 5; Length 22;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Gy 2 TCCCTCACCTCCAGGCACAT 21
 |||||
 Db 1 TTCACTACCTCCATCCACAT 20

RESULT 6

ACH62883
 ID ACH62883 standard; DNA, 25 BP.
 XX
 AC ACH62883;
 XX
 DT 17-OCT-2003 (first entry)
 XX

DE DNA target sequence #12019 useful in array for genetic analyses.

XX Gene expression analysis; array; hybridisation; genetic variation;
 XX tag-labelled compound; gene family; in situ hybridisation;
 KW library screening; Southern hybridisation; northern hybridisation;
 KW dot-blot hybridisation; gene sequence; mutation detection;
 KW target sequence; probe; PCR; primer; ss.
 XX

OS Unidentified.
 XX
 PN US2003082596-A1.
 XX
 PD 01-MAY-2003.
 XX
 PF 08-AUG-2002; 2002US-00215112.
 XX
 PR 08-AUG-2001; 2001US-0311040P.
 XX
 PA (MITT/) MITTMANN M.
 XX
 PI Mittmann M;
 XX
 DR WPI; 2003-576608/54.
 XX
 PT New probe array useful e.g. for monitoring gene expression levels, for
 PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
 PT comprises multiple nucleic acid probes.
 XX
 PS Claim 1; SEQ ID NO 12019; 9pp; English.

CC The present invention relates to nucleic acid sequences that are
 CC complementary to particular genes, and can be used as probes for a
 CC variety of analyses such as gene expression analysis. Each probe
 CC comprises 9 or more consecutive nucleotides from at least one of 14936
 CC nucleotide sequences defined in the patent, or their perfect sense match,
 CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
 CC The probes may be used in an array comprising at least 10 distinct
 CC nucleic acid probes. The array is useful in monitoring gene expression
 CC levels by hybridisation to a DNA library, in analysing genetic
 CC variations, and in hybridising tag-labelled compounds. The probes are
 CC useful for identifying family members of a gene. The probes are also
 CC useful in situ hybridisations, in screening cDNA or genomic libraries
 CC (or derived subclones) for additional clones containing segments of DNA
 CC that have been previously isolated and sequenced. In Southern, northern,
 CC or dot-blot hybridisation of genomic DNA to identify or detect the
 CC sequence of any gene or detect specific mutations in any gene, and in
 CC mapping the 5' termini of mRNA molecules by primer extensions. The
 CC nucleic acid sequences of the invention are also useful as PCR primers.
 CC The invention provides a large collection of nucleic acid sequences
 CC complementary to particular genes with a wide range of analytical uses.
 CC ACH50865-ACH65260 represent the target sequences of the invention. Note:
 CC The sequence data for this patent was obtained in electronic format
 CC directly from the USPTO web site at seqdata.uspto.gov/patidbidentry.htm1
 XX
 XX
 SQ Sequence 25 BP; 6 A; 11 C; 4 G; 4 T; 0 U; 0 Other;

Query Match 69.1%; Score 15.2; DB 9; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Gy 3 CCCCTCACCTCCAGGCACATG 22
 |||||
 Db 1 CCTTACGCTCCACACCATG 20

RESULT 7

ACH63009
 ID ACH63009 standard; DNA, 25 BP.
 XX
 AC ACH63009;
 XX
 DT 17-OCT-2003 (first entry)
 XX

DE DNA target sequence #12145 useful in array for genetic analyses.

XX Gene expression analysis; array; hybridisation; genetic variation;
 XX tag-labelled compound; gene family; in situ hybridisation;
 KW library screening; Southern hybridisation; northern hybridisation;
 KW dot-blot hybridisation; gene sequence; mutation detection;
 KW target sequence; probe; PCR; primer; ss.
 XX

OS Unidentified.
 XX
 PN US2003082596-A1.
 XX
 PD 01-MAY-2003.
 XX
 PF 08-AUG-2002; 2002US-00215112.
 XX
 PR 08-AUG-2001; 2001US-0311040P.
 XX
 PA (MITT/) MITTMANN M.
 XX
 PI Miltmann M;
 XX
 DR WPI; 2003-576608/54.
 XX
 PT New probe array useful e.g. for monitoring gene expression levels, for
 PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
 PT comprising multiple nucleic acid probes.
 XX
 PS Claim 1; SEQ ID NO 12145; 9pp; English.
 XX
 CC The present invention relates to nucleic acid sequences that are
 CC complementary to particular genes, and can be used as probes for a
 CC variety of analyses such as gene expression analysis. Each probe
 CC comprises 9 or more consecutive nucleotides from at least one of 14936
 CC nucleotide sequences defined in the patent, or their perfect sense match,
 CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
 CC The probe may be used in an array comprising at least 10 distinct
 CC nucleic acid probes. The array is useful in monitoring gene expression
 CC levels by hybridisation to a DNA library, in analysing genetic
 CC variations, and in hybridising tag-labelled compounds. The probes are
 CC useful for identifying family members of a gene. The probes are also
 CC useful in situ hybridisations, in screening cDNA or genomic libraries
 CC (or derived subclones) for additional clones containing segments of DNA
 CC that have been previously isolated and sequenced, in Southern, northern,
 CC or dot-blot hybridisation of genomic DNA to identify or detect the
 CC sequence of any gene or detect specific mutations in any gene, and in
 CC mapping the 5' termini of mRNA molecules by primer extensions. The
 CC nucleic acid sequences of the invention are also useful as PCR primers.
 CC The invention provides a large collection of nucleic acid sequences
 CC complementary to particular genes with a wide range of analytical uses.
 CC ACH50865-ACH55260 represent the target sequences of the invention. Note:
 CC The sequence data for this patent was obtained in electronic format
 CC directly from the USPTO web site at seqdata.uspto.gov/patid/patidntry.html
 XX
 SQ Sequence 25 BP; 5 A; 11 C; 4 G; 5 T; 0 U; 0 Other;
 Query Match 69.1%; Score 15.2; DB 9; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 CCTCTACCTCGAGCCACATG 22
 Db 1 CCTTCAGCTCATCCACATG 20

RESULT 8
 AAV55821
 ID AAV55821 standard; DNA; 24 BP.
 XX
 AC AAV55821;
 XX
 DT 27-AUG-2003 (revised)
 DT 18-NOV-1998 (first entry)
 XX
 DE Multimerisation of minimal motifs using primer ZGY2.
 XX
 KM Fusion protein; stabilising polypeptide; proteolytic degradation;
 KM resistance; half-life; autoimmune disease; inflammation; nitro drug;
 KM IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
 KM nitroductase protein; enzyme therapy; produg therapy; protease;
 KM cancer; pathological condition; minimal motif; PCR primer; ss.

XX
 OS Synthetic.
 OS Human herpesvirus 4.
 XX
 PN MO9822577-A1.
 XX
 PD 28-MAY-1998.
 XX
 PF 17-NOV-1997; 97WO-IB001508.
 XX
 PR 15-NOV-1996; 96US-0030986P.
 PR 25-JUN-1997; 97US-0048945P.
 XX
 PA (MASU/) MASUCCI M G.
 XX
 PI Masucci MG;
 XX
 DR WPI; 1998-312463/27.
 XX
 PT New fusion proteins resistant to proteolytic degradation - comprising a
 PT core protein with a stabilising polypeptide comprising a peptide sequence
 PT containing glycine repeats.
 XX
 PS Disclosure; Page 72; 120pp; English.
 XX
 CC Sequences shown in AAV55812 to AAV55827 represent primers used in the
 CC course of the invention for the multimerisation of minimal motifs. The
 CC invention provides a method for increasing the resistance of a core
 CC protein to proteolytic degradation that comprises linking or inserting
 CC onto or into the core protein a stabilising polypeptide of formula
 CC [(Gly)X(Glyb)Y(Glyc)Z]n where Glya, Glyb, Glyc are 1-6 sequential Gly
 CC residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met, Phe, Pro or Thr
 CC and n can be anything between 1-66. X, Y and Z need not be identical from
 CC n repeat to n repeat. Alternatively a nucleic acid encoding a stabilising
 CC polypeptide can be linked onto or inserted into a nucleic acid encoding a
 CC core protein. The fusion proteins of the invention are more resistant to
 CC degradation by proteases and, thus, have a longer half-life than the
 CC unfused core protein. The products can be used for treating autoimmune
 CC diseases, cancer and inflammation. In particular the core protein may be
 CC an IkappaB regulator protein for the treatment of inflammatory bowel
 CC disease, or a nitroductase protein which can activate nitro drugs in
 CC enzyme/prodrug therapy to treat cancer or other pathological conditions.
 CC The fusion proteins can also be used in diagnostic methods such as in
 CC vivo imaging. (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 24 BP; 5 A; 13 C; 2 G; 4 T; 0 U; 0 Other;
 Query Match 66.4%; Score 14.6; DB 2; Length 24;
 Best Local Similarity 81.0%; Pred. No. 2.3e+04;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CTCCTCAGCTCGAGCCACAT 21
 Db 4 CACCCGACCTCCAGCTCAT 24

RESULT 9
 ADY00944/c
 ID ADY00944 standard; DNA; 29 BP.
 XX
 AC ADY00944;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE PCR primer 170 used to amplify human GALE region SNP DNA.
 XX
 KM SNP detection; breast tumor; endocrine disease;
 KM gynecology and obstetrics; neoplasia; cytostatic; metastasis;
 KM gene therapy; RNA interference; UDP-galactose 4-epimerase; GALE;
 KM UDP-glucose 4-epimerase; lysophospholipase; Ht014; LOC148902; LYPLA2; ss;
 KM PCR; primer.
 XX
 OS Homo sapiens.

CC genes designed from repetitive peptide sequences, such as glycoproteins
CC (including the peptide sequences of gum arabic glycoprotein (GAGP)).
CC ABL51730 to ABL51849 and ABB78401 to ABB78544 represent sequences used in
CC the exemplification of the present invention

SO Sequence 30 BP; 6 A, 19 C, 0 G, 5 T, 0 U, 0 Other;

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCACTCCACCCAT 21
Db 5 CACCTTCACTCCACCCCAT 25

RESULT 12

ABL51739
ID ABL51739 standard; DNA; 30 BP.

AC ABL51739;

DT 09-JUL-2002 (first entry)

XX Hydroxyproline-rich glycoprotein (HRGP) related linker SEQ ID NO:38.

XX Plant; Gum arabic glycoprotein; GAGP; hydroxyproline-rich glycoprotein;
XX HRGP; repetitive proline-rich protein; RRP; arabinogalactan protein;
XX AGP; plant gum; PCR primer; linker; ss.

OS Acacia senegal.
XX Synthetic.

PN WO200178503-A2.

PD 25-OCT-2001.

PF 12-APR-2001; 2001WO-US012336.

PR 12-APR-2000; 2000US-00547693.

PA (UYOH-) UNIV OHIO.

PI Kieliszewski MJ;

DR WPI; 2002-041307/05.

XX Nucleic acids and proteins useful for producing hydroxy-proline rich
PT glycoproteins in plants.

PS Example 2; Page 53; 326pp; English.

XX The present invention describes synthetic genes encoding plant gums and
CC other hydroxyproline (Hyp)-rich glycoproteins (HRGP) and the nucleic
CC acids that encode them. The nucleic acids, proteins and methods from the
CC present invention may be used to produce HRGP, repetitive proline-rich
CC proteins (RRP) and arabinogalactan-proteins (AGP) in plants via
CC recombinant methodologies. Also described is the expression of synthetic
CC genes designed from repetitive peptide sequences, such as glycoproteins
CC (including the peptide sequences of gum arabic glycoprotein (GAGP)).
CC ABL51730 to ABL51849 and ABB78401 to ABB78544 represent sequences used in
CC the exemplification of the present invention

SO Sequence 30 BP; 6 A, 19 C, 0 G, 5 T, 0 U, 0 Other;

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCACTCCACCCAT 21
Db 5 CACCTTCACTCCACCCCAT 25

RESULT 13

ADU09143
ID ADU09143 standard; DNA; 30 BP.

AC ADU09143;

DT 27-JAN-2005 (first entry)

XX Ser-Hyp4 gene cassette oligonucleotide #7.

XX Gum arabic glycoprotein; GAGP; Hydroxyproline-rich glycoprotein; HRGP;
XX repetitive proline-rich protein; RRP; arabinogalactan-protein; AGP;
XX plant gum; elastin; extensin; thickener; emulsifier; stabiliser;
XX surfactant; flavour fixative; food industry; cosmetic; ss.

OS Synthetic.

PN WO2004094590-A2.

PD 04-NOV-2004.

PF 13-APR-2004; 2004WO-US011174.

PR 16-APR-2003; 2003US-00418032.

PA (UYOH-) UNIV OHIO.

PI Kieliszewski MJ;

DR WPI; 2004-784900/77.

XX Plant gums (especially gum arabic) useful in the food and cosmetics
PT industries are produced by expression of synthetic genes designed from
PT the repetitive peptide sequences of plant gum glycoproteins.

PS Example 2; SEQ ID NO 38; 179pp; English.

XX The invention relates to an isolated polypeptide (I) comprising
CC alternating rigid and non-rigid modules, where the rigid modules are
CC directed by contiguous hydroxyproline residues and the non-rigid modules
CC are directed by non-contiguous hydroxyproline residues. The polypeptides
CC are synthetic gums based on motifs and repeats found in Gum arabic
CC glycoprotein (GAGP), Hydroxyproline-rich glycoproteins (HRGP), repetitive
CC proline-rich proteins (RRP) and arabinogalactan-proteins (AGP). Also
CC included are a polypeptide (II) (comprising a first elastin module
CC flanked on the N-terminal side by a first extensin module having a first
CC crosslinkage motif, and a second elastin module flanked on the C-terminal
CC side by a second extensin module having a second crosslinkage motif) and
CC a nucleic acid encoding a polypeptide variant of Lycopodium esculentum
CC arabinogalactan protein (Le-AGP-1). In polypeptide (I), the rigid module
CC comprises at least two contiguous hydroxyproline residues and is
CC arabinosylated. Polypeptide (II) further comprises a glycoprotein having a
CC crosslinkage motif. The crosslinkage motif comprises tyrosine, preferably
CC the crosslinkage motif is Val-Tyr-Tyr. The tyrosine forms an
CC intramolecular or intermolecular crosslink. In polypeptide (II), the
CC first elastin module is repeated six times and the second elastin module
CC is repeated three times. The first and/or second crosslinkage motifs
CC preferably comprise Val-Tyr-Tyr and contiguous hydroxyproline residues.
CC (II) further comprises a central stretch of rigid arabinosylated Ser-O-O-
CC O-O (where O is hydroxyproline, appearing as ADU09108) repeats flanked on
CC either side by the first and the second elastin modules. The Le-AGP-1
CC variant encoded by lacks a glycosylphosphatidylinositol anchor signal
CC sequence and an internal lysine rich region. The variant encoded by the
CC nucleic acid further comprises an enhanced green fluorescent protein
CC fusion glycoprotein. The nucleic acid is expressed in Nicotiana tabacum.
CC The polypeptide (I) is useful as a thickener, emulsifier, stabiliser,
CC surfactant, or flavour fixative in the food industry. It is also useful
CC in cosmetics. The polypeptide (I) has an anomalously low viscosity. The
CC present sequence encodes a cassette peptide comprising one or more
CC consensus motifs or repeat units derived from GAGP, AGP, RRP or HRGP
XX used to make a polypeptide of the invention.

Mon Mar 27 08:53:26 2006

us-10-798-652-6.sz1m30.png

Page 9

Job time : 375.5 secs

This Page Blank (uspto)

QY	4	CCTCACCTCCAGACC	17
Db	15	CCTCACCTCCAGACC	2
RESULT 2	D21049	25 bp	mRNA linear EST 30-JUL-1996
LOCUS	D21049	HMGSS02034 Human promyelocyte Homo sapiens cDNA clone mp0338 3', mRNA sequence.	
DEFINITION	D21049	GI:504869	EST.
ACCESSION	D21049	GI:504869	EST.
VERSION	D21049.1	GI:504869	EST.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;
Best Local Similarity	80.0%;	Pred. No. 3.5e+05;	
Matches	16;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	1	CTCCTCACCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3	AZ980990/c	22 bp	DNA linear GSS 27-APR-2001
LOCUS	AZ980990/c	22 bp	DNA linear GSS 27-APR-2001
DEFINITION	2M0258F17P Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0258F17 F, genomic survey sequence.		
ACCESSION	AZ980990		
VERSION	AZ980990.1	GI:13852217	GSS.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;
Best Local Similarity	80.0%;	Pred. No. 3.5e+05;	
Matches	16;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	1	CTCCTCACCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3	AZ980990	22 bp	DNA linear GSS 27-APR-2001
LOCUS	2M0258F17P Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0258F17 F, genomic survey sequence.		
ACCESSION	AZ980990		
VERSION	AZ980990.1	GI:13852217	GSS.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;
Best Local Similarity	80.0%;	Pred. No. 3.5e+05;	
Matches	16;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	1	CTCCTCACCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3	AZ980990	22 bp	DNA linear GSS 27-APR-2001
LOCUS	2M0258F17P Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0258F17 F, genomic survey sequence.		
ACCESSION	AZ980990		
VERSION	AZ980990.1	GI:13852217	GSS.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;
Best Local Similarity	80.0%;	Pred. No. 3.5e+05;	
Matches	16;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	1	CTCCTCACCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3	AZ980990	22 bp	DNA linear GSS 27-APR-2001
LOCUS	2M0258F17P Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0258F17 F, genomic survey sequence.		
ACCESSION	AZ980990		
VERSION	AZ980990.1	GI:13852217	GSS.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;
Best Local Similarity	80.0%;	Pred. No. 3.5e+05;	
Matches	16;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	1	CTCCTCACCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3	AZ980990	22 bp	DNA linear GSS 27-APR-2001
LOCUS	2M0258F17P Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0258F17 F, genomic survey sequence.		
ACCESSION	AZ980990		
VERSION	AZ980990.1	GI:13852217	GSS.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	61.8%	Score 13.6;	DB 8; Length 25;</

University of Utah
Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0258 row: F column: 17
Seq primer: CGTTGTTAAACGACCGCACT
Class: plasmid ends
High quality sequence stop: 22.
Location/Qualifiers
1. .22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U96CM0258F17"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/clone_1lb="Mouse 10kb plasmid U96CM library"
/note="Vector: FMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adapter oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g14732114[gp]Arl29072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 60.04; Score 13.2; DB 9; Length 22;
Best Local Similarity 83.34; Pred. No.5.2e+05;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CCTGCACTCCAGCCACA 20
||| ||||| |||||
Db 22 CCCCACCTCCACACACA 5

RESULT 4
A2655652/c 24 bp DNA linear GSS 14-DEC-2000
LOCUS A2655652
DEFINITION IM0530D1R Mouse 10kb plasmid U96CM library Mus musculus genomic
clone U96CM0530D1R, genomic survey sequence.
ACCESSION A2655652
VERSION A2655652.1 GI:11792798
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 24)
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausen,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished. (2000)
Contact: Robert B. Weiss

JOURNAL
COMMENT

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0530 row: D column: 13

Seq primer: CACACAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 24.

FEATURES

source

Location/Qualifiers

1..24

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="GenBank:57BL/6J"

/db_xref="taxon:10090"

/clone="UUCG1M0530D13"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource"

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114[g14732114], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match

Best Local Similarity 59.1%; Score 13; DB 9; Length 24;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

2 TCCCTCAGCTCCAGCCACATG 22

23 TACCTGACATCCACCCACAG 3

RESULT 5

AA864650/c

LOCUS

DEFINITION

AA864650 28 bp mRNA linear EST 13-MAY-1998
oh37b09.g1 NCI CGAP Kid6 Homo sapiens cDNA IMAGE:1459961.3,
similar to SW:ATPO_EOVIN P13620 ATP SYNTHASE D CHAIN, MITOCHONDRIAL
; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominoidea; Homo.

NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc.

DNA Sequencing by: Washington University Genome Sequencing Center

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bdip/image/image.html

Insert Length: 952 Std Error: 0.00

Seq primer: -40ml3 fwd. RT from Amersham

High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers

1..28

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/clone="IMAGE:1459961"

/sex="mixed"

/tissue_type="kidney tumor"

/lab_host="SOLR (kanamycin resistant)"

/clone_lib="NCI CGAP Kid6"

/note="Organ: kidney; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATTCGACGAG 3' 3' adaptor sequence: 5'
CTCAGCTTTTCTTTTCTTTT 3' Average insert size: 1.0 kb."

ORIGIN

Query Match

Best Local Similarity 59.1%; Score 13; DB 1; Length 28;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3 CCCCTCAGCTCCAG 15

27 CCCCTCAGCTCCAG 15

RESULT 6

AZ498825/c

LOCUS

DEFINITION

AZ498825 29 bp DNA linear GSS 05-OCT-2000
IM0336J1F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG1M0336J1 F, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Muridae; Murinae; Mus.

1 (bases 1 to 29)

Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,

Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von

Niederhausen,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0336 row: J column: 14

Seq primer: CGTTGTAAACGACGCGCACT

Class: plasmid ends

High quality sequence stop: 29.

Location/Qualifiers

1..29

`/organism="Mus musculus"`
`/mol_type="genomic DNA"`
`/strain="C57BL/6J"`
`/db_xref="taxon:10090"`
`/clone="UGC1M03363J4"`
`/sex="Male"`
`/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"`
`/clone_lib="Mouse 10kb plasmid UGC1M library"`
`/note="Vector: PMD24ny1. Purified genomic DNA from M.`
`musculus C57BL/6J (male). was obtained from the Jackson`
`Laboratory Mouse DNA Resource`
`(http://www.jax.org/resources/documents/dnares/). The DNA`
`was hydrodynamically sheared by repeated passage through a`
`0.005 inch orifice at constant velocity. The sheared DNA`
`was blunt end-repaired with T4 DNA polymerase and T4`
`polynucleotide kinase. Adaptor oligonucleotides were`
`ligated to the blunt ends in high molar excess. The`
`adaptored DNA was purified and size-selected for a 9.5 to`
`10.5 kb range using preparative agarose gel`
`electrophoresis. Vector DNA was prepared from a derivative`
`of PMD2 (G114733114[9]AF12972.1), a copy-number`
`inducible derivative of plasmid R1. The vector was ligated`
`with adaptors complementary to the insert adaptors and`
`purified. The sheared, adaptored mouse DNA was annealed to`
`adaptored vector DNA, and transformed into`
`chemically-competent E. coli XL10-Gold (Stratagene) cells`
`and selected for ampicillin resistance."`

ORIGIN

Query Match	58.2%	Score 12.8;	DB 9;	Length 29;
Best Local Similarity	87.5%;	Pred. No. 7.7e+05;		
Matches 14;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

Qy 2 TCCCTCACCCTCCAGCC 17
|||
Db 24 TCCCTCAGATCCAGCC 9

RESULT	7
AZ778032/c	
LOCUS	AZ778032
DEFINITION	22 bp DNA linear GSS 16-FEB-2001 2K001A11F Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUCG2M001A11 F, genomic survey sequence.

ACCESSION	A2778032	
VERSION	A2778032.1	GI:12907248
KEYWORDS	GSS.	
SOURCE	Mus musculus	(house mouse)

ORGANISM
Mus musculus
Mys myosculus
Bakaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sclurgonathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

TITLE	Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid insert
COMMENT	Unpublished (2000) Contact: Robert B. Weiss

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

```

FEATURES
Email: ddm@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0013 row: A column: 11
Seq primer: CGGTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stcp: 22.
Location/Qualifiers

```

Source

1. 22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0013A11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_id="Mouse 10kb plasmid UGGCM library"
/note="Vector: PWD4291. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrolytically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative of
pMD22 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match	57.3%	Score 12.6;	DB 9;	Length 22;
Best Local Similarity	78.9%	Pred. No. 9.2e+05;		
Matches 15; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	CTCCCTACCTCCAGCCAC	19
Db	22	CCCCACACCTCTACCCAC	4

LOCUS	DEFINITION	23 bp	DNA	1linear	GSS 16-FEB-2003
A2785027					
2M0028H03	Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC2M0028H03 R, genomic survey sequence.				

ACCESSION	AZ785027
VERSION	AZ785027.1
KEYWORDS	GSS.
SOURCE	Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE AUTHORS

TITLE	Mouse, whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid inserts
COMMENT	Unpublished (2000) Contact: Robert B. Weiss

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0028 row: H column: 03
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 23.

FEATURES
SOURCE

Location/Qualifiers
1. .23
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC2M0028H03"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UUC1M library"
/note="Vector: PMD42nv, Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passages through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 57.3%; Score 12.6; DB 9; Length 23;
Best Local Similarity 78.9%; Pred. No. 9.2e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCAGCCAC 19
Db 19 CCCCTCAGCTGACAGAC 1

RESULT 9 25 bp DNA linear GSS 20-OCT-2003
CG730897/c 119130F05.2BL_x1 1119 - RescueMu Grid AA Zea mays genomic, genomic
LOCUS
DEFINITION
survey sequence.

ACCESSION CG730897
VERSION CG730897.1 GI:37773389
KEYWORDS
SOURCE
ORGANISM

Zea mays
Zea mays
Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 25)
Walbot.V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 119130 row: F column: 05
Class: transposon-tagged.

FEATURES
SOURCE

Location/Qualifiers
1. .25
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"

/db_xref="taxon:4577"
/issue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="1119 - RescueMu Grid AA"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI, Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 57.3%; Score 12.6; DB 10; Length 25;
Best Local Similarity 78.9%; Pred. No. 9.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCAGCCAC 19
Db 24 CTTATCGCTCTCGCCAC 6

RESULT 10 19 bp DNA linear GSS 02-OCT-2000
A2381798/c 1M0138G01R Mouse 10kb plasmid UUC1M library Mus musculus genomic
LOCUS
DEFINITION
clone UUC1M0138G01 R, genomic survey sequence.

ACCESSION A2381798

VERSION A2381798.1 GI:10495498
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus

Bakaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE

TITILE

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0138 row: G column: 01
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
SOURCE

Location/Qualifiers
1. .19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0138G01"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UUC1M library"
/note="Vector: PMD42nv, Purified genomic DNA from M.

/clone_11b="Mouse 10kb plasmid UGCG1M library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 55.5%; Score 12.2; DB 9; Length 27;
Best Local Similarity 82.4%; Pred. No. 1.4e+06;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4 CCTCAGCTCCAGCCACA 20
Db 25 CCTCAGCTCCAGCCACA 9

RESULT 13

LOCUS A2599021 23 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0414F08F Mouse 10kb plasmid UGCG1M library Mus musculus genomic
clone UGCG1M0414F08 F, genomic survey sequence.

ACCESSION A2599021
VERSION A2599021.1 GI:11721211
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,B., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0414 row: F column: 08
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 23.

Location/Qualifiers

FEATURES

Source 1..23
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0414F08"
/sex="Male"

/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_11b="Mouse 10kb plasmid UGCG1M library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 54.5%; Score 12; DB 9; Length 23;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CTCCTCAGCTCCAGCCACA 20
Db 21 CTCCTCAGCTCCAGCCACA 2

RESULT 14

LOCUS A2762096 24 bp DNA linear GSS 16-FEB-2001
DEFINITION 1M0556J18 Mouse 10kb plasmid UGCG1M library Mus musculus genomic
clone UGCG1M0556J18 R, genomic survey sequence.

ACCESSION A2762096
VERSION A2762096.1 GI:12871740
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 24)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,B., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0556 row: J column: 18
Seq primer: CACACAGCAACACGATGACC
Class: plasmid ends
High quality sequence stop: 24.

Location/Qualifiers

FEATURES

Source 1..24
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0556J18"

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCM library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid p1. The vector was ligated
with adaptor complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

```

Query Match          54.5%; Score 12; DB 9; Length 24;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CTCCTCAGCTCCAGCCACA 20
    |||||
DB 24 CCCCCCCCCGCGCCGCCACA 5

```

RESULT 15

LOCUS

```

A1416870      25 bp      mRNA      linear      EST 23-UTL-2004
bail9b03.x1 Gm-c1005 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1005-102 3 similar to TR:Q39599 Q39599 EXTENSIN. ;, mRNA
sequence.

```

ACCESSION

VERSION

```

A1416870
A1416870.1 GI:4260374

```

KEYWORDS

SOURCE

ORGANISM

```

Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 25)

```

REFERENCE

AUTHORS

```

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M.,
Bowers,X., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

```

TITLE

JOURNAL

COMMENT

```

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

```

```

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. GENOME SYSTEMS
CLONE ID: Gm-c1005-102 Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand this clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163, email: info@biogeneticservices.com)
Seq primer: T7 ET from Amersham
High quality sequence stop: 1
POLYA=No.

```

FEATURES

source

```

Location/Qualifiers
1..25
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1005-102"
/lab_host="XL10-Gold"
/clone_lib="Gm-c1005"
/note="Vector: pBluescript II XE; Site 1: EcoRI; Site 2:
XhoI. This library was constructed by Dr. Randy C.
Shoemaker and Dr. John Erpelding, USDA-ARS Agronomy
Department, G401 Agronomy Hall, Iowa State University,
 Ames, IA 50011-1010 Phone: 515-294-6233 FAX: 515-294-2299
email: rcshoe@iastate.edu"

```

ORIGIN

```

Query Match          54.5%; Score 12; DB 1; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.7e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CTCCTCAGCTCCAGCCACA 20
    |||||
DB 6 CCCCCCTCAGCCGCCA 25

```

```

Search completed: March 25, 2006, 15:28:17
Job time : 3045.5 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 12:58:33 ; Search time 129.5 Seconds
(without alignments)
301.980 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22
Sequence: 1 ctccctcaccctccagccacatg 22

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 1026780

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/pdata/1/ina/1_COMB.seq:*
2: /cgn2_6/pdata/1/ina/5_COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/H_COMB.seq:*
6: /cgn2_6/pdata/1/ina/PCRTUS_COMB.seq:*
7: /cgn2_6/pdata/1/ina/PP_COMB.seq:*
8: /cgn2_6/pdata/1/ina/RG_COMB.seq:*
9: /cgn2_6/pdata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.6	66.4	24	2	US-08-529-190B-16
2	14.6	66.4	30	3	US-09-119-507B-10
3	14.6	66.4	30	3	US-09-119-507B-38
4	14.6	66.4	30	3	US-08-897-556A-10
5	14.6	66.4	30	3	US-08-897-556A-38
6	14.6	66.4	30	3	US-09-547-693-10
7	14.6	66.4	30	3	US-09-547-693-38
8	14.2	64.5	24	2	US-08-529-190B-14
9	14.2	64.5	24	3	US-07-792-600-8
10	14.2	64.5	24	3	US-09-157-021-8
11	14.2	64.5	24	3	US-09-156-842-8
12	14.2	64.5	24	3	US-09-591-514-8
13	14.2	64.5	24	3	US-09-887-145-6
14	13.8	62.7	20	3	US-09-674-824-5
15	13.6	61.8	20	3	US-09-487-368A-216
16	13.6	61.8	20	3	US-09-658-679A-52
17	13.6	61.8	20	3	US-09-629-644A-216
18	13.6	61.8	20	3	US-09-629-644A-216
19	13.6	61.8	24	2	US-08-529-190B-4
20	13.6	61.8	24	2	US-08-529-190B-7
21	13.6	61.8	25	3	US-09-396-196G-40878
22	13.6	61.8	25	3	US-09-396-196G-101880
23	13.6	61.8	27	2	US-08-460-890A-19
24	13.6	61.8	27	2	US-08-460-890A-22

C 25	13.6	61.8	27	3	US-08-167-641C-19	Sequence 19, Appl
C 26	13.6	61.8	27	3	US-08-167-641C-22	Sequence 22, Appl
C 27	13.6	61.8	27	3	US-08-460-971A-19	Sequence 19, Appl
C 28	13.6	61.8	27	3	US-08-460-971A-22	Sequence 22, Appl
C 29	13.6	61.8	27	3	US-08-462-040-19	Sequence 19, Appl
C 30	13.6	61.8	27	3	US-08-462-040-22	Sequence 22, Appl
C 31	13.4	60.9	25	3	US-09-396-196G-59434	Sequence 59434, A
C 32	13.2	60.0	19	3	US-09-358-972-221	Sequence 221, App
C 33	13.2	60.0	19	3	US-09-406-064-28	Sequence 28, Appl
C 34	13.2	60.0	19	3	US-09-430-615-58	Sequence 58, Appl
C 35	13.2	60.0	19	3	US-09-406-065-62	Sequence 62, Appl
C 36	13.2	60.0	19	3	US-09-383-316-78	Sequence 78, Appl
C 37	13.2	60.0	19	3	US-09-788-847-28	Sequence 28, Appl
C 38	13.2	60.0	19	3	US-09-790-417-221	Sequence 221, App
C 39	13.2	60.0	22	9	5219727-31	Patent No. 5219727
C 40	13.2	60.0	24	2	US-08-529-190B-2	Sequence 2, Appl
C 41	13.2	60.0	24	2	US-08-529-190B-5	Sequence 5, Appl
C 42	13.2	60.0	24	2	US-08-463-978-2	Sequence 2, Appl
C 43	13.2	60.0	24	2	US-08-463-978-5	Sequence 5, Appl
C 44	13.2	60.0	24	6	PCT-US94-05467-2	Sequence 2, Appl
C 45	13.2	60.0	24	6	PCT-US94-05467-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-529-190B-16
Sequence 16, Application US/08529190B
Patent No. 5833991
GENERAL INFORMATION:
APPLICANT: Masucci, Maria G.
TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
TITLE OF INVENTION: CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: One Financial Center
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,190B
FILING DATE: 15-SEP-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE9501324-9
FILING DATE: 10-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/522,595
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen A
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/53015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-529-190B-16

Query Match 66.4%; Score 14.6; DB 2; Length 24;
Best Local Similarity 81.0%; Pred. No. 3.8e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 4 CACCCGACCTCCAGCTCCAT 24

RESULT 2

US-09-119-507B-10
; Sequence 10, Application US/09119507B
; Patent No. 6548642
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia J.
; TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums
; FILE REFERENCE: OHU-03417
; CURRENT APPLICATION NUMBER: US/09/119,507B
; CURRENT FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-119-507B-10

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 3

US-09-119-507B-38
; Sequence 38, Application US/09119507B
; Patent No. 6548642
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia J.
; TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums
; FILE REFERENCE: OHU-03417
; CURRENT APPLICATION NUMBER: US/09/119,507B
; CURRENT FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-119-507B-38

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 4

US-08-897-556A-10
; Sequence 10, Application US/08897556A
; Patent No. 657062
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA J.

; TITLE OF INVENTION: SYNTHETIC GENES FOR PLANT GUMS AND OTHER
; TITLE OF INVENTION: HYDROXYPROLINE-RICH GLYCOPROTEINS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,556A
; FILING DATE: 21-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OHU-02908
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-897-556A-10

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 5

US-08-897-556A-38
; Sequence 38, Application US/08897556A
; Patent No. 657062
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA J.
; TITLE OF INVENTION: SYNTHETIC GENES FOR PLANT GUMS AND OTHER
; TITLE OF INVENTION: HYDROXYPROLINE-RICH GLYCOPROTEINS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,556A
; FILING DATE: 21-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837

```

; REFERENCE/DOCKET NUMBER: OHU-02908
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-897-556A-38

Query Match
Best Local Similarity 66.4%; Score 14.6; DB 3; Length 30;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCACCTCCAGCCACAT 21
Db 5 CACCTTCACCTCCAGCCCAT 25

RESULT 6
US-09-547-693-10
; Sequence 10, Application US/09547693
; Patent No. 6639050
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia
; TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich
; FILE REFERENCE: OHU-04089
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 236
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 10
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Synthetic
US-09-547-693-10

Query Match
Best Local Similarity 66.4%; Score 14.6; DB 3; Length 30;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCACCTCCAGCCACAT 21
Db 5 CACCTTCACCTCCAGCCCAT 25

RESULT 7
US-09-547-693-38
; Sequence 38, Application US/09547693
; Patent No. 6639050
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia
; TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich
; FILE REFERENCE: OHU-04089
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 236
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 38
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: misc_feature
```

```

; OTHER INFORMATION: Synthetic
US-09-547-693-38

Query Match
Best Local Similarity 66.4%; Score 14.6; DB 3; Length 30;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCACCTCCAGCCACAT 21
Db 5 CACCTTCACCTCCAGCCCAT 25

RESULT 8
US-08-529-190B-14/C
; Sequence 14, Application US/08529190B
; Patent No. 583391
; GENERAL INFORMATION:
; APPLICANT: Masucci, Maria G.
; TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,190B
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE9501324-9
; FILING DATE: 10-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/522,595
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen A
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3255/53015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-529-190B-14

Query Match
Best Local Similarity 64.5%; Score 14.2; DB 2; Length 24;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CCTTCACCTCCAGCCACAT 21
Db 23 CCGCAGCTCCAGCTCCAT 5

RESULT 9
US-07-792-600-8
; Sequence 8, Application US/07792600
; Patent No. 6008045
; GENERAL INFORMATION:
```

APPLICANT: COPELAND, WILLIAM C.
APPLICANT: WANG, TERESA S.-F.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: TEMPLATE-DEPENDENT ENZYMATIC SYNTHESIS OF NUCLEIC ACID
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Peter G. Carroll
STREET: 220 Montgomery Street, Suite 710
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/792,600
FILING DATE: 19911115
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: STDU-00097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear

US-07-792-600-8

Query Match 64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTCACCCTCCAGCCACATG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 1 CTTCACTCCAGCCAGGTG 19

RESULT 10
US-09-157-021-8
Sequence 8, Application US/09157021A
Patent No. 6100023
GENERAL INFORMATION:
APPLICANT: Copeland, William C.
APPLICANT: Wang, Teresa S. F.
TITLE OF INVENTION: Drug Design Assay
FILE REFERENCE: STDU-03484
CURRENT APPLICATION NUMBER: US/09/157,021A
CURRENT FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: 07/792,600
EARLIER FILING DATE: 1991-11-15
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-157-021-8

Query Match 64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTCACCCTCCAGCCACATG 22

DB 1 | | | | | | | | | | | | | | | | | | | | | |
1 CTTCACTCCAGCCAGGTG 19

RESULT 11
US-09-156-842-8
Sequence 8, Application US/09156842A
Patent No. 6103473
GENERAL INFORMATION:
APPLICANT: Copeland, William C.
APPLICANT: Wang, Teresa S. F.
TITLE OF INVENTION: Drug Screening
FILE REFERENCE: STDU-03485
CURRENT APPLICATION NUMBER: US/09/156,842A
CURRENT FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: 07/792,600
EARLIER FILING DATE: 1991-11-15
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-156-842-8

Query Match 64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTCACCCTCCAGCCACATG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 1 CTTCACTCCAGCCAGGTG 19

RESULT 12
US-09-591-514-8
Sequence 8, Application US/09591514
Patent No. 6670161
GENERAL INFORMATION:
APPLICANT: Copeland, William C.
APPLICANT: Wang, Teresa S. F.
TITLE OF INVENTION: Drug Design Assay
FILE REFERENCE: STDU-03484
CURRENT APPLICATION NUMBER: US/09/591,514
CURRENT FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: US/09/157,021
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 07/792,600
PRIOR FILING DATE: 1991-11-15
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-591-514-8

Query Match 64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTCACCCTCCAGCCACATG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 1 CTTCACTCCAGCCAGGTG 19

RESULT 13
US-09-887-145-6/c
Sequence 6, Application US/09887145


```
Patent No. 6780641
GENERAL INFORMATION:
APPLICANT: Kim, Seung U
TITLE OF INVENTION: immortalized human microglia
cell and continuous cell line
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: David Praehker, Esq.
STREET: P.O. Box 5387
CITY: Magnolia
STATE: Massachusetts
COUNTRY: USA
ZIP: 01930
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
COMPUTER: Dell PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Microsoft Word version 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/887,145
FILING DATE: 22-Jun-2001
CLASSIFICATION: Unknown
ATTORNEY/AGENT INFORMATION:
NAME: David Praehker, Esq.
REGISTRATION NUMBER: 29,693
REFERENCE/DOCKET NUMBER: UBC-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (978) 525-3794
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-887-145-6

Query Match          63.6%; Score 14; DB 3; Length 24;
Best Local Similarity 77.3%; Pred. No. 6.8e+03;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCAGCCATG 22
Db 22 CACCTCAGCCATCAATG 1

RESULT 14
US-09-674-824-5/c
Sequence 5, Application US/09674824
Patent No. 6890732
GENERAL INFORMATION:
APPLICANT: Lotz, et al
TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH CODE FOR ENZYMES DERIVED FROM WHEAT
FILE REFERENCE: AGR 1998/M 205; FLH514413-3848
CURRENT APPLICATION NUMBER: US/09/574,824
CURRENT FILING DATE: 2000-02-16
PRIOR APPLICATION NUMBER: PCT/EP99/03156
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: DE 198 20607.0
PRIOR FILING DATE: 1998-05-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence, primer
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(20)
OTHER INFORMATION: primer
US-09-674-824-5
```

```
Query Match          62.7%; Score 13.8; DB 3; Length 20;
Best Local Similarity 88.2%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CTCACCTCCAGCCACA 20
Db 17 CCTCAGCTGAGCCACA 1

RESULT 15
US-09-487-368A-216
Sequence 216, Application US/09487368A
Patent No. 6261840
GENERAL INFORMATION:
APPLICANT: Lex M. Cowseart
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF PTP1B EXPRESSION
FILE REFERENCE: RTS-0093
CURRENT APPLICATION NUMBER: US/09/487,368A
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 240
SEQ ID NO 216
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-487-368A-216

Query Match          61.8%; Score 13.6; DB 3; Length 20;
Best Local Similarity 80.0%; Pred. No. 9.8e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCAGCCACA 20
Db 1 CTGCCCCAGCAGCCAGCCACA 20
```

Search completed: March 25, 2006, 13:03:00
Job time : 130.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using BW model

Run on: March 25, 2006, 13:59:41 ; Search time 660.5 Seconds

(without alignments)
275.437 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22

Sequence: 1 CTCCTCACCCTCCAGCCACATG 22

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 10535742

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	8	US-10-798-652-6
2	20.4	92.7	22	8	US-10-798-652-7
3	15.8	71.8	25	8	US-10-719-900-593282
4	15.8	71.8	25	9	US-10-956-157-252303
5	15.6	70.9	25	9	US-10-485-048-86
6	15.6	70.9	25	10	US-11-036-317-179644
7	15.6	70.9	25	10	US-11-036-317-250940
8	15.6	70.9	25	10	US-11-036-317-253919
9	15.6	70.9	25	10	US-11-036-317-268615
10	15.6	70.9	25	10	US-11-036-317-362146
11	15.6	70.9	25	10	US-11-036-317-391184
12	15.4	70.0	25	10	US-11-036-317-579816
13	15.2	69.1	25	5	US-10-215-112-12019
14	15.2	69.1	25	5	US-10-215-112-12145
15	15.2	69.1	25	10	US-11-036-317-278200
16	15.2	69.1	25	10	US-11-036-317-386177
17	14.8	67.3	25	8	US-10-719-900-217185
18	14.8	67.3	25	8	US-10-719-900-217186
19	14.8	67.3	25	9	US-10-956-157-132785
20	14.6	66.4	25	7	US-10-719-956-459914
21	14.6	66.4	25	10	US-11-036-317-280075
22	14.6	66.4	25	10	US-11-036-317-360952
23	14.6	66.4	25	10	US-11-036-317-469308

24	14.6	66.4	25	10	US-11-036-317-516411	Sequence 516411,
C 25	14.6	66.4	25	10	US-11-036-317-897472	Sequence 897472,
C 26	14.6	66.4	25	10	US-11-036-317-927798	Sequence 927798,
C 27	14.6	66.4	25	10	US-11-036-317-980808	Sequence 980808,
C 28	14.6	66.4	25	10	US-11-060-756-206000	Sequence 206000,
29	14.6	66.4	25	10	US-11-060-756-20618	Sequence 20618, A
30	14.6	66.4	30	6	US-10-437-708-10	Sequence 10, Appl
31	14.6	66.4	30	6	US-10-437-708-38	Sequence 38, Appl
32	14.6	66.4	30	6	US-10-395-402-10	Sequence 10, Appl
33	14.6	66.4	30	6	US-10-395-402-38	Sequence 38, Appl
34	14.6	66.4	30	8	US-10-257-199-10	Sequence 10, Appl
35	14.6	66.4	30	8	US-10-257-199-38	Sequence 38, Appl
36	14.6	66.4	30	9	US-10-418-032-10	Sequence 10, Appl
37	14.6	66.4	30	9	US-10-418-032-38	Sequence 38, Appl
38	14.2	64.5	25	7	US-10-681-773-98598	Sequence 98598, A
39	14.2	64.5	25	7	US-10-681-773-98599	Sequence 98599, A
C 40	14.2	64.5	25	8	US-10-719-900-221421	Sequence 221421,
C 41	14.2	64.5	25	8	US-10-719-900-693283	Sequence 693283,
42	14.2	64.5	25	9	US-10-956-157-70581	Sequence 70581, A
43	14.2	64.5	25	9	US-10-956-157-70582	Sequence 70582, A
44	14.2	64.5	25	9	US-10-956-157-70583	Sequence 70583, A
45	14.2	64.5	25	9	US-10-956-157-70584	Sequence 70584, A

ALIGNMENTS

RESULT 1
US-10-798-652-6
Sequence 6, Application US/10798652
Publication No. US20040219582A1
GENERAL INFORMATION:
APPLICANT: Yongjun Guo
TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
FILE REFERENCE: FGF-3 GENE AND METHODS OF USE THEREOF
CURRENT APPLICATION NUMBER: US/10/798,652
CURRENT FILING DATE: 2004-03-11
PRIOR APPLICATION NUMBER: 60/455,698
PRIOR FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe
US-10-798-652-6

Query Match 100.0%; Score 22; DB 8; Length 22;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCCAGCCACATG 22
Db 1 CTCCTCACCCTCCAGCCACATG 22

RESULT 2
US-10-798-652-7
Sequence 7, Application US/10798652
Publication No. US20040219582A1
GENERAL INFORMATION:
APPLICANT: Yongjun Guo
TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
FILE REFERENCE: FGF-3 GENE AND METHODS OF USE THEREOF
CURRENT APPLICATION NUMBER: US/10/798,652
CURRENT FILING DATE: 2004-03-11
PRIOR APPLICATION NUMBER: 60/455,698
PRIOR FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 7
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe
US-10-798-652-7

Query Match 92.7%; Score 20.4; DB 8; Length 22;
Best Local Similarity 95.5%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 1 CTCCTCGCCTCCAGCCACATG 22

RESULT 3
US-10-719-900-693282/c
Sequence 693282, Application US/10719900
Publication No. US20050026164A1
GENERAL INFORMATION:
APPLICANT: Xue Mei Zhou
TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
FILE REFERENCE: 3528.1
CURRENT APPLICATION NUMBER: US/10/719,900
CURRENT FILING DATE: 2003-11-20
PRIOR APPLICATION NUMBER: 60/427,808
PRIOR FILING DATE: 2002-11-20
NUMBER OF SEQ ID NOS: 982914
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 693282
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-10-719-900-693282

Query Match 71.8%; Score 15.8; DB 8; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCCTCACTCCAGCCACA 20
DB 22 TCCCTCACTCCAGCCACA 4

RESULT 4
US-10-956-157-252303/c
Sequence 252303, Application US/10956157
Publication No. US20050118625A1
GENERAL INFORMATION:
APPLICANT: Myeth
APPLICANT: Mounth, William
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
FILE REFERENCE: 031896-043000 (AM 101081)
CURRENT APPLICATION NUMBER: US/10/956,157
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 319805
SOFTWARE: Patent in version 3.2
SEQ ID NO 252303
LENGTH: 25
TYPE: DNA
ORGANISM: Probe Sequence
US-10-956-157-252303

Query Match 71.8%; Score 15.8; DB 9; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCAC 19
DB 1 CTCCTCACTCCAGCCAC 19

DB 24 CTCCTCACTCCAGCCAC 6

RESULT 5
US-10-485-048-86/c
Sequence 86, Application US/10485048
Publication No. US20050058657A1
GENERAL INFORMATION:
APPLICANT: ERTL, Peter Franz
APPLICANT: TITE, John Philip
APPLICANT: VAN MEY, Catherine Anne
APPLICANT: VOSS, Gerald
TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
FILE REFERENCE: B45279
CURRENT APPLICATION NUMBER: US/10/485,048
CURRENT FILING DATE: 2004-01-26
PRIOR APPLICATION NUMBER: PCT/EP02/08343
PRIOR FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: GB 0118367.2
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 98
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 86
LENGTH: 25
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer (Figure 21)
US-10-485-048-86

Query Match 70.9%; Score 15.6; DB 9; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 22 CTCCTCACTCCAGCCACATG 1

RESULT 6
US-11-036-317-179644/c
Sequence 179644, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Blume, John
APPLICANT: Williams, Alan
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 179644
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-179644

Query Match 70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 23 CCCACTCCCTCCAGCCACAG 2

RESULT 7
US-11-036-317-250940
Sequence 250940, Application US/11036317

```
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 250940
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-250940

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 2 CTCCTCCCATGAGCCACCTG 23

RESULT 8
US-11-036-317-253919/c
/ Sequence 253919, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 253919
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-253919

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 22 CCCACTCCCTCCAGCCACAG 1

RESULT 9
US-11-036-317-268615
/ Sequence 268615, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 268615
```

```
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-268615

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 1 CTCCTCCCATGAGCCACCTG 22

RESULT 10
US-11-036-317-362146
/ Sequence 362146, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 362146
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-362146

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 3 CTCCTCCCATGAGCCACCTG 24

RESULT 11
US-11-036-317-391184
/ Sequence 391184, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 391184
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-391184

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 4 CTCCTCCCATGAGCCACCTG 25
```

```
RESULT 12
US-11-036-317-579816/c
; Sequence 579816, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO: 579816
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-579816

Query Match          70.0%; Score 15.4; DB 10; Length 25;
Best Local Similarity 94.1%; Pred. No. 6.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2  TCCTCAGCTCCAGCCCA 18
Db      19 TCCTCAGCTCCAGCCCA 3

RESULT 13
US-10-215-112-12019
; Sequence 12019, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; PRIOR FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12019
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-12019

Query Match          69.1%; Score 15.2; DB 5; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      3  CCCTCAGCTCCAGCCCATG 22
Db      1  CCTTCAGTCCACACCATG 20

RESULT 14
US-10-215-112-12145
; Sequence 12145, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; PRIOR FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12145
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-12145

Query Match          69.1%; Score 15.2; DB 5; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      3  CCCTCAGCTCCAGCCCATG 22
Db      1  CCTTCAGTCCATCCACCATG 20

RESULT 15
US-11-036-317-278200/c
; Sequence 278200, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO: 278200
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-278200

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  CTCCTCAGCTCCAGCCACA 20
Db      21 CCCTCAGCTCCAGCCACA 2

Search completed: March 25, 2006, 15:50:31
Job time : 661.5 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 14:11:31 ; Search time 315.5 Seconds
(without alignments)
277.964 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22
Sequence: 1 ccccccacccacgacacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9258654 seqs, 1993127192 residues

Total number of hits satisfying chosen parameters: 13041728

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*

1: /SIDS5/pcodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/pcodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/pcodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/pcodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/pcodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/pcodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/pcodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /SIDS5/pcodata/1/pubpna/US10_NEW_PUB.seq2:*
9: /SIDS5/pcodata/1/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/pcodata/1/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/pcodata/1/pubpna/US11_NEW_PUB.seq2:*
12: /SIDS5/pcodata/1/pubpna/US11_NEW_PUB.seq3:*
13: /SIDS5/pcodata/1/pubpna/US11_NEW_PUB.seq4:*
14: /SIDS5/pcodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/pcodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	79.1	20	8	US-10-310-914A-613434 Sequence 613434
2	16.8	76.4	24	8	US-10-310-914A-233742 Sequence 233742
3	16.4	74.5	19	8	US-10-310-914A-668834 Sequence 668834
4	16.4	74.5	20	8	US-10-310-914A-317759 Sequence 317759
5	16.4	74.5	20	8	US-10-310-914A-377554 Sequence 377554
6	16.2	73.6	25	8	US-10-310-914A-91862 Sequence 91862
7	16.2	73.6	25	8	US-10-310-914A-91863 Sequence 91863
8	15.8	71.8	20	8	US-10-310-914A-448406 Sequence 448406
9	15.8	71.8	21	8	US-10-310-914A-724597 Sequence 724597
10	15.8	71.8	23	8	US-10-310-914A-448407 Sequence 448407
11	15.8	71.8	23	8	US-10-310-914A-1002544 Sequence 1002544
12	15.6	70.9	24	8	US-10-310-914A-180307 Sequence 180307
13	15.6	70.9	24	8	US-10-310-914A-66290 Sequence 66290
14	15.4	70.0	19	12	US-11-101-244-196424 Sequence 196424
15	15.4	70.0	19	13	US-11-083-784-196424 Sequence 196424
16	15.4	70.0	23	8	US-10-310-914A-228828 Sequence 228828
17	15.2	69.1	20	8	US-10-310-914A-290068 Sequence 290068
18	15.2	69.1	24	8	US-10-310-914A-662777 Sequence 662777

C 19	14.8	67.3	20	8	US-10-310-914A-701328 Sequence 701328
C 20	14.8	67.3	21	8	US-10-310-914A-1108644 Sequence 1108644
C 21	14.8	67.3	23	8	US-10-310-914A-233763 Sequence 233763
C 22	14.8	67.3	23	8	US-10-310-914A-586537 Sequence 586537
C 23	14.8	67.3	23	8	US-10-310-914A-771600 Sequence 771600
C 24	14.8	67.3	24	8	US-10-310-914A-306671 Sequence 306671
C 25	14.8	67.3	25	14	US-11-121-849-31947 Sequence 31947
C 26	14.8	67.3	25	14	US-11-136-527-165393 Sequence 165393
C 27	14.8	67.3	25	14	US-11-136-527-165398 Sequence 165398
C 28	14.8	67.3	25	14	US-11-136-527-165399 Sequence 165399
C 29	14.8	67.3	25	14	US-11-136-527-165399 Sequence 165399
C 30	14.8	67.3	25	14	US-11-136-527-165406 Sequence 165406
C 31	14.8	67.3	26	8	US-10-310-914A-233764 Sequence 233764
C 32	14.8	67.3	26	8	US-10-310-914A-367833 Sequence 367833
C 33	14.6	66.4	23	8	US-10-310-914A-601024 Sequence 601024
C 34	14.6	66.4	24	8	US-10-310-914A-872970 Sequence 872970
C 35	14.6	66.4	24	8	US-10-310-914A-1231595 Sequence 1231595
C 36	14.6	66.4	25	14	US-11-121-849-631263 Sequence 631263
C 37	14.6	66.4	29	8	US-10-857-780-2372 Sequence 2372
C 38	14.4	65.5	18	8	US-10-310-914A-668835 Sequence 668835
C 39	14.4	65.5	18	8	US-10-310-914A-730873 Sequence 730873
C 40	14.4	65.5	18	8	US-10-310-914A-849102 Sequence 849102
C 41	14.4	65.5	18	8	US-10-310-914A-1127878 Sequence 1127878
C 42	14.4	65.5	19	12	US-11-101-244-464103 Sequence 464103
C 43	14.4	65.5	19	12	US-11-101-244-769382 Sequence 769382
C 44	14.4	65.5	19	12	US-11-101-244-773064 Sequence 773064
C 45	14.4	65.5	19	12	US-11-101-244-1035410 Sequence 1035410

ALIGNMENTS

RESULT 1
US-10-310-914A-613434
; Sequence 613434, Application US/10310914A
; Publication No. US2006000322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Biointernationally detectable group of novel regulatory genes an
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310, 914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 613434
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
; US-10-310-914A-613434

Query Match 79.1%; Score 17.4; DB 8; Length 20;
Best Local Similarity 78.9%; Pred. No. 4.9e+02;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCTGACTCCAGCCACAT 21
DB 2 CCGGACCCGACGACACAU 20

RESULT 2
US-10-310-914A-233742
; Sequence 233742, Application US/10310914A
; Publication No. US2006000322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Biointernationally detectable group of novel regulatory genes an
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310, 914A
; CURRENT FILING DATE: 2002-12-06

NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 233742
LENGTH: 24
TYPE: RNA
ORGANISM: Human
US-10-310-914A-233742

Query Match 76.4%; Score 16.8; DB 8; Length 24;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCCTCAGCTCCAGCCACATG 22
DB 1 CCCCAACCTCCAGCCACATG 20

RESULT 3
US-10-310-914A-668834/c
Sequence 668834, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 668834
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-668834

Query Match 74.5%; Score 16.4; DB 8; Length 19;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCTCAGCTCCAGCCACA 20
DB 18 CCCCACTCCAGCCACA 1

RESULT 4
US-10-310-914A-317759
Sequence 317759, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 317759
LENGTH: 20
TYPE: RNA
ORGANISM: Human
US-10-310-914A-317759

Query Match 74.5%; Score 16.4; DB 8; Length 20;
Best Local Similarity 77.8%; Pred. No. 1.3e+03;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 CCTCAGCTCCAGCCACAT 21
DB 3 CCUACCCUCCGCCACAU 20

RESULT 5
US-10-310-914A-377554/c
Sequence 377554, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 377554
LENGTH: 20
TYPE: RNA
ORGANISM: Human
US-10-310-914A-377554

Query Match 74.5%; Score 16.4; DB 8; Length 20;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCA 18
DB 18 CCCCTCAGCTCCAGCCA 1

RESULT 6
US-10-310-914A-91862/c
Sequence 91862, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 91862
LENGTH: 25
TYPE: RNA
ORGANISM: Human
US-10-310-914A-91862

Query Match 73.6%; Score 16.2; DB 8; Length 25;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCTCAGCTCCAGCCATG 22
DB 23 TCCTTAAGCTCCAGCTCTG 3

RESULT 7
US-10-310-914A-91863/c
Sequence 91863, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402


```
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 91863
 ; LENGTH: 25
 ; TYPE: RNA
 ; ORGANISM: Human
US-10-310-914A-91863
```

```
Query Match
Beat Local Similarity 73.6%; Score 16.2; DB 8; Length 25;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 TCCTCAGCTCCAGCCACATG 22
DB 23 TCCTTAAGCTCCAGCTCCTG 3
```

```
RESULT 8
US-10-310-914A-448406
 ; Sequence 448406, Application US/10310914A
 ; Publication No. US20060003322A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bentwich, Isaac
 ; APPLICANT: Shlier, Kvuzae
 ; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
 ; FILE REFERENCE: 06087.0200.CPUS01
 ; CURRENT APPLICATION NUMBER: US/10/310,914A
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 1388402
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 448406
 ; LENGTH: 20
 ; TYPE: RNA
 ; ORGANISM: Human
US-10-310-914A-448406
```

```
Query Match
Beat Local Similarity 71.8%; Score 15.8; DB 8; Length 20;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCAGCTCCAGCCAC 19
DB 1 CUCCGACACUCCAGCCAC 19
```

```
RESULT 9
US-10-310-914A-724597/c
 ; Sequence 724597, Application US/10310914A
 ; Publication No. US20060003322A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bentwich, Isaac
 ; APPLICANT: Shlier, Kvuzae
 ; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
 ; FILE REFERENCE: 06087.0200.CPUS01
 ; CURRENT APPLICATION NUMBER: US/10/310,914A
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 1388402
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 724597
 ; LENGTH: 21
 ; TYPE: RNA
 ; ORGANISM: Human
US-10-310-914A-724597
```

```
Query Match
Beat Local Similarity 71.8%; Score 15.8; DB 8; Length 21;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 TCCTCAGCTCCAGCCACA 20
DB 19 TCCTCATCTCTGCCACA 1
```

```
RESULT 10
US-10-310-914A-448407
 ; Sequence 448407, Application US/10310914A
 ; Publication No. US20060003322A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bentwich, Isaac
 ; APPLICANT: Shlier, Kvuzae
 ; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes an
 ; FILE REFERENCE: 06087.0200.CPUS01
 ; CURRENT APPLICATION NUMBER: US/10/310,914A
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 1388402
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 448407
 ; LENGTH: 23
 ; TYPE: RNA
 ; ORGANISM: Human
US-10-310-914A-448407
```

```
Query Match
Beat Local Similarity 71.8%; Score 15.8; DB 8; Length 23;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCAGCTCCAGCCAC 19
DB 1 CUCCGACACUCCAGCCAC 19
```

```
RESULT 11
US-10-310-914A-1002544
 ; Sequence 1002544, Application US/10310914A
 ; Publication No. US20060003322A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bentwich, Isaac
 ; APPLICANT: Shlier, Kvuzae
 ; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
 ; FILE REFERENCE: 06087.0200.CPUS01
 ; CURRENT APPLICATION NUMBER: US/10/310,914A
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 1388402
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 1002544
 ; LENGTH: 23
 ; TYPE: RNA
 ; ORGANISM: Human
US-10-310-914A-1002544
```

```
Query Match
Beat Local Similarity 71.8%; Score 15.8; DB 8; Length 23;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCAGCTCCAGCCAC 19
DB 4 CUCCGACACUCCAGCCAC 22
```

```
RESULT 12
US-10-310-914A-180307/c
 ; Sequence 180307, Application US/10310914A
 ; Publication No. US20060003322A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bentwich, Isaac
 ; APPLICANT: Shlier, Kvuzae
 ; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
 ; FILE REFERENCE: 06087.0200.CPUS01
 ; CURRENT APPLICATION NUMBER: US/10/310,914A
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 1388402
 ; SOFTWARE: PatentIn version 3.3
```

SEQ ID NO 180307
LENGTH: 24
TYPE: RNA
ORGANISM: Human
US-10-310-914A-180307

Query Match 71.8%; Score 15.8; DB 8; Length 24;
Best Local Similarity 89.5%; Pred. No. 2.4e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTCAGCTCCAGCCACATG 22
DB 24 CCTCACTCCAGCCACCTG 6

RESULT 13

US-10-310-914A-66290
Sequence 66290, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazak
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 66290
LENGTH: 24
TYPE: RNA
ORGANISM: Human
US-10-310-914A-66290

Query Match 70.9%; Score 15.6; DB 8; Length 24;
Best Local Similarity 72.7%; Pred. No. 2.9e+03;
Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACATG 22
DB 2 CACCCACCCACCCUAGACACAU 23

RESULT 14

US-11-101-244-196424/C
Sequence 196424, Application US/11101244
Publication No. US20050246794A1
GENERAL INFORMATION:
APPLICANT: Dharmoon, Inc.
APPLICANT: Khvorova, Anastasia
APPLICANT: Reynolds, Angela
APPLICANT: Leake, Devin
APPLICANT: Marshall, William
APPLICANT: Scaringe, Stephen
TITLE OF INVENTION: Functional and Hyperfunctional siRNA
FILE REFERENCE: 13499US
CURRENT APPLICATION NUMBER: US/11/101,244
CURRENT FILING DATE: 2005-04-07
PRIOR APPLICATION NUMBER: 60/502,050
PRIOR FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: 60/426,137
PRIOR FILING DATE: 2002-11-14
NUMBER OF SEQ ID NOS: 1591911
SOFTWARE: Proprietary
SEQ ID NO 196424
LENGTH: 19
TYPE: RNA
ORGANISM: Homo sapiens
US-11-101-244-196424

Query Match 70.0%; Score 15.4; DB 12; Length 19;
Best Local Similarity 94.1%; Pred. No. 3.4e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTCCAGCC 17
DB 17 CTCCTCAGCTCCAGCC 1

RESULT 15
US-11-083-784-196424/C
Sequence 196424, Application US/11083784
Publication No. US20050245475A1
GENERAL INFORMATION:
APPLICANT: Dharmoon, Inc.
APPLICANT: Khvorova, Anastasia
APPLICANT: Reynolds, Angela
APPLICANT: Leake, Devin
APPLICANT: Marshall, William
APPLICANT: Scaringe, Stephen
TITLE OF INVENTION: Functional and Hyperfunctional siRNA
FILE REFERENCE: 13499US
CURRENT APPLICATION NUMBER: US/11/083,784
CURRENT FILING DATE: 2005-03-18
PRIOR APPLICATION NUMBER: US/10/714,333
PRIOR FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 60/502,050
PRIOR FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: 60/426,137
PRIOR FILING DATE: 2002-11-14
NUMBER OF SEQ ID NOS: 1591911
SOFTWARE: Proprietary
SEQ ID NO 196424
LENGTH: 19
TYPE: RNA
ORGANISM: Homo sapiens
US-11-083-784-196424

Query Match 70.0%; Score 15.4; DB 13; Length 19;
Best Local Similarity 94.1%; Pred. No. 3.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCC 17
DB 17 CTCCTCAGCTCCAGCC 1

Search completed: March 25, 2006, 14:54:04
Job time : 315.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 13:42:31 ; Search time 959 Seconds

(without alignments)
1304.021 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22

Sequence: 1 ctcctcgctccagccacacg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 1641224

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

1: gb_da: *
2: gb_in: *
3: gb_env: *
4: gb_ov: *
5: gb_ov: *
6: gb_ov: *
7: gb_ov: *
8: gb_ov: *
9: gb_ov: *
10: gb_ov: *
11: gb_ov: *
12: gb_ov: *
13: gb_ov: *
14: gb_ov: *
15: gb_ov: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	64.5	30	6	A79795 Sequence 3
2	14	63.6	25	6	AX701087 Sequence
3	13.8	62.7	20	6	CS109308 Sequence
4	13.8	62.7	24	6	AX44501 Sequence
5	13.6	61.8	22	6	BD086268 Vector. 8
6	13.6	61.8	22	6	AX003404 Sequence
7	13.6	61.8	27	6	AR091427 Sequence
8	13.6	61.8	27	6	AR091430 Sequence
9	13.6	61.8	27	6	AR125632 Sequence
10	13.6	61.8	27	6	AR125635 Sequence
11	13.6	61.8	27	6	AR125632 Sequence
12	13.6	61.8	27	6	AR125632 Sequence
13	13.6	61.8	27	6	AR125632 Sequence
14	13.6	61.8	27	6	AR125632 Sequence
15	13.6	61.8	27	6	AR125632 Sequence
16	13.6	61.8	27	6	AR125632 Sequence
17	13.6	61.8	27	6	AR125632 Sequence
18	13.6	61.8	27	6	AR125632 Sequence

19	13	59.1	30	6	AR340063 Sequence
20	13	59.1	30	6	AR412123 Sequence
21	13	59.1	30	6	AR412132 Sequence
22	12.8	58.2	17	6	AR379426 Sequence
23	12.8	58.2	20	6	AX598365 Sequence
24	12.8	58.2	21	6	BD174200 Sequence
25	12.8	58.2	21	6	BD185155 Sequence
26	12.8	58.2	21	6	CS012518 Sequence
27	12.8	58.2	22	6	AX817728 Sequence
28	12.6	57.3	19	6	CS095314 Sequence
29	12.6	57.3	19	6	CS095413 Sequence
30	12.6	57.3	22	6	BD249866 Sequence
31	12.6	57.3	22	6	AX033501 Sequence
32	12.6	57.3	24	6	A57517 Sequence 9
33	12.6	57.3	24	6	A57536 Sequence 28
34	12.6	57.3	24	6	AR052983 Sequence
35	12.6	57.3	24	6	AR096578 Sequence
36	12.6	57.3	24	6	AR105954 Sequence
37	12.6	57.3	24	6	AR442798 Sequence
38	12.6	57.3	24	6	AX290630 Sequence
39	12.6	57.3	25	6	BD174183 Sequence
40	12.6	57.3	25	6	BD185138 Cell diff
41	12.6	57.3	25	6	AX501644 Sequence
42	12.6	57.3	25	6	AX501645 Sequence
43	12.6	57.3	25	6	AX501646 Sequence
44	12.6	57.3	25	6	AX501647 Sequence
45	12.6	57.3	25	6	AX501648 Sequence

ALIGNMENTS

RESULT 1

LOCUS A79795 Sequence 3 from Patent WO9714795. 30 bp DNA linear PAT 20-OCT-1999

DEFINITION A79795

ACCESSION A79795

VERSION A79795.1 GI:6092713

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

ORIGIN

Query Match

Best local Similarity

Matches

16; Conservative

0; Mismatches

3; Indels

0; Gaps

0; Score 14.2; DB 6; Length 30;

Pred. No. 1.4e+05;

4 CCTGCGCTCCAGCCATG 22

|||||

|||||

|||||

|||||

|||||

|||||

|||||

ORGANISM	synthetic construct other sequences; artificial sequences.					
REFERENCE	1 Ertl,P.F., Tite,J.P., van Wely,C.A. and Voss,G. Vaccine comprising gp120 and nef and/or tat for the immunisation against hiv					
AUTHORS	Patent: WO 0301334-A 86 13-FEB-2003;					
TITLE	Glaxosmithkline Biologicals S.A. (BE) ; GLAXO GROUP LIMITED (GB)					
JOURNAL						
FEATURES	Location/Qualifiers 1..25 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="primer (figure 21)"					
SOURCE						
ORIGIN						
Query Match	63.6%;	Score 14;	DB 6;	Length 25;		
Best Local Similarity	77.3%;	Pred.No. 1.7e+05;				
Matches	17;	Conservative	0;	Mismatches	5;	Indels
OY	1 CTCCTCGCCCTTCGACGCACATG	22				
Db	22 CTTTCCCACTCCTCGCACATG	1				
RESULT 3						
CS109308						
LOCUS	CS109308	20 bp	DNA	linear	PAT 22-JUN-2005	
DEFINITION	Sequence 7 from Patent WO2005052140.					
ACCESSION	CS109308					
VERSION	CS109308.1	GI:68148122				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 Shi,S., Seo,B.M. and Mlura,M. Multipotent postnatal stem cells from human periodontal ligament and uses thereof Patent: WO 2005052140-A 7 09-JUN-2005;					
AUTHORS	Department of Health and Human Services (US)					
TITLE	Location/Qualifiers 1..20 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"					
JOURNAL						
FEATURES						
SOURCE						
ORIGIN						
Query Match	62.7%;	Score 13.8;	DB 6;	Length 20;		
Best Local Similarity	88.2%;	Pred.No. 2.1e+05;				
Matches	15;	Conservative	0;	Mismatches	2;	Indels
OY	5 CTCGGCTCGACGCACAT	21				
Db	1 CTGGCCTCGACTACAT	17				
RESULT 4						
AX44501/c						
LOCUS	AX44501	24 bp	DNA	linear	PAT 03-JUL-2002	
DEFINITION	Sequence 956 from Patent WO216649.					
ACCESSION	AX44501					
VERSION	AX44501.1	GI:21691779				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 Gunderson,K. Probes and decoder oligonucleotides Patent: WO 0216649-A 956 28-FEB-2002;					
AUTHORS	Department of Health and Human Services (US)					
TITLE	Location/Qualifiers 1..20 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"					
JOURNAL						
FEATURES						
SOURCE						
ORIGIN						

FEATURES	Illumina, Inc. (US)
source	Location/Qualifiers 1..24 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Computer Generated Probe Sequence."
ORIGIN	
Query Match	62.7%; Score 13.8; DB 6; Length 24;
Best Local Similarity	88.2%; Pred. No. 2.1e+05;
Matches 15; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Oy	5 CTGCGCTTCAGCCACAT 21
Db	21 CTCGCGTTCAGACACAT 5
RESULT 5	
LOCUS	BD086268 22 bp DNA linear PAT 27-AUG-2002
DEFINITION	Vector.
ACCESSION	BD086268
VERSION	BD086268.1 GI:22631878
KEYWORDS	JP 2001525168-A/21. synthetic construct synthetic construct other sequences; artificial sequences. 1 (bases 1 to 22) Shen,S., Schedl,A. and Harmar,A.J. Vector
SOURCE	Patent: JP 2001525168-A 21 11-DEC-2001; MEDICAL RESEARCH COUNCIL OS Artificial Sequence PN JP 2001525168-A/21 PD 11-DEC-2001 PF 27-NOV-1998 JP 2000523236 PR 28-NOV-1997 GB 9725311.6,28-NOV-1997 GB 9725313.2 PR 20-MAR-1998 GB 9806072.6,05-NOV-1998 GB 9824275.3 PI SANBING SHEN, ANDREAS SCHEDL,ANTHONY JOHN HARMAR PC C12N15/09,C12N15/00 CC Description of Artificial Sequence: Primer FH Key Location/Qualifiers FT source 1..22 Location/Qualifiers 1..22 /organism="Artificial Sequence". Location/Qualifiers 1..22 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
FEATURES	
source	Location/Qualifiers 1..22 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
ORIGIN	
Query Match	61.8%; Score 13.6; DB 6; Length 22;
Best Local Similarity	80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Oy	2 TCCTCGCTTCAGCCACAT 21
Db	1 TTCATCACCTCATCCACAT 20
RESULT 6	
LOCUS	AX003404 22 bp DNA linear PAT 24-AUG-2000
DEFINITION	Sequence 21 from Patent WO9298449.
ACCESSION	AX003404
VERSION	AX003404.1 GI:9927208
KEYWORDS	. synthetic construct synthetic construct other sequences; artificial sequences.
SOURCE	
ORGANISM	
REFERENCE	1 Schedl,A. and Harmar,A.J.
AUTHORS	

TITLE VectorB
JOURNAL Patent: WO 9928449-A 21 10-JUN-1999;
SCHEDL ANDREAS (DE); HARMAR ANTHONY JOHN (GB)
FEATURES
SOURCE Location/Qualifiers
1..22
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"

ORIGIN
Query Match 61.8%; Score 13.6; DB 6; Length 22;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TCCCTGCGCTCCAGCCACAT 21
Db 1 TTCAATCACCCTCCATCCACAT 20

RESULT 7
AR091427/c
LOCUS AR091427 27 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 19 from patent US 5994109.
ACCESSION AR091427
VERSION AR091427.1 GI:10018182
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter system and methods of use
JOURNAL Patent: US 5994109-A 19 30-NOV-1999;
FEATURES
SOURCE Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 8
AR091430/c
LOCUS AR091430 27 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 22 from patent US 5994109.
ACCESSION AR091430
VERSION AR091430.1 GI:10018185
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter system and methods of use
JOURNAL Patent: US 5994109-A 22 30-NOV-1999;
FEATURES
SOURCE Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 9
AR125632/c
LOCUS AR125632 27 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 19 from patent US 617554.
ACCESSION AR125632
VERSION AR125632.1 GI:14111694
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter systems
JOURNAL Patent: US 617554-A 19 23-JAN-2001;
FEATURES
SOURCE Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 10
AR125635/c
LOCUS AR125635 27 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 22 from patent US 617554.
ACCESSION AR125635
VERSION AR125635.1 GI:14111697
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter systems
JOURNAL Patent: US 617554-A 22 23-JAN-2001;
FEATURES
SOURCE Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 11
AX732524
LOCUS AX732524 17 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 4158 from Patent WO03025175.
ACCESSION AX732524
VERSION AX732524.1 GI:30511867
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
AUTHORS Telerman, A., Amson, R. and Tuijinder, M.
TITLE Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or virus resistance and their use as
medicines
JOURNAL Patent: WO 03025175-A 4158 27-MAR-2003;
Molecular Engines Laboratories (FR)
FEATURES
source
1..17
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 59.1%; Score 13; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CCTCCAGCCACAT 21
| | | | | | | | | |
Db 5 CCTCCAGCCACAT 17

RESULT 12
AR073296 23 bp DNA linear PAT 28-AUG-2000
LOCUS
DEFINITION Sequence 10 from patent US 5948892.
ACCESSION AR073296
VERSION AR073296.1 GI:10000059
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 23)
AUTHORS Wahl, R.C.
TITLE Analogs of macrophage stimulating protein
JOURNAL Patent: US 5948892-A 10 07-SEP-1999;
FEATURES
source
1..23
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 59.1%; Score 13; DB 6; Length 23;
Best Local Similarity 76.2%; Pred. No. 4.8e+05;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCCCTGCCTCCAGCCACATG 22
| | | | | | | | | |
Db 2 TGGTTCTCTCTCCGCGCATATG 22

RESULT 13
A57518 24 bp DNA linear PAT 03-MAR-1998
LOCUS
DEFINITION Sequence 10 from Patent WO9632483.
ACCESSION A57518
VERSION A57518.1 GI:3713376
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE unclassified sequences.
1
AUTHORS Masucci, M.G.
TITLE IMMUNE-EVADING PROTEINS
JOURNAL Patent: WO 9632483-A 10 17-OCT-1996;
MASUCCI MARIA GRAZIA (SE)
COMMENT Other publication AU 5284296 961030.
FEATURES
source
1..24
/organism="unidentified"

ORIGIN
Query Match 59.1%; Score 13; DB 6; Length 24;
Best Local Similarity 76.2%; Pred. No. 4.8e+05;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCCTGCCTCCAGCCACAT 21
| | | | | | | | | |
Db 4 CACCCGCACTCCAGCTCCAT 24

RESULT 14
AR052984 24 bp DNA linear PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 16 from patent US 5833991.
ACCESSION AR052984
VERSION AR052984.1 GI:5977846
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 24)
AUTHORS Masucci, M.G.
TITLE Glycine-containing sequences conferring invisibility to the immune
system
JOURNAL Patent: US 5833991-A 16 10-NOV-1998;
FEATURES
source
1..24
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 59.1%; Score 13; DB 6; Length 24;
Best Local Similarity 76.2%; Pred. No. 4.8e+05;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCCTGCCTCCAGCCACAT 21
| | | | | | | | | |
Db 4 CACCCGCACTCCAGCTCCAT 24

RESULT 15
CQ876350 25 bp DNA linear PAT 04-OCT-2004
LOCUS
DEFINITION Sequence 200 from Patent WO2004065583.
ACCESSION CQ876350
VERSION CQ876350.1 GI:53789954
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
1
AUTHORS Cobleigh, M.A., Shak, S., Baker, J.B. and Cronin, M.T.
TITLE Gene expression markers for breast cancer prognosis
JOURNAL Patent: WO 2004065583-A 200 05-AUG-2004;
Genomic Health, Inc. (US); Rush University Medical Center (US)
FEATURES
source
1..25
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32640"
/note="probe"

ORIGIN
Query Match 59.1%; Score 13; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.8e+05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CTGCGCTCCAGCC 17
| | | | | | | | | |
Db 12 CTGCGCTCCAGCC 24

Search completed: March 25, 2006, 14:43:30
Job time : 960 secs

This Page Blank (uspto)

XX Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
PS Claim 20; SEQ ID NO 7, 15pp; English.

The invention relates to an isolated nucleic acid molecule comprising the upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3 gene sequence (Adb797901) with a single nucleotide polymorphism (SNP) at position 69. Also included are an isolated nucleic acid molecule comprising a sequence complementary to Adb797901, a vector comprising Adb797901 operably linked to a reporter gene, a host cell comprising the vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating a nucleic acid sample from the mammal, and determining whether a cytosine or thymine is present at position 69 of Adb797901), a kit for performing the method (comprising a first oligonucleotide probe which anneals specifically with the target portion of the mammal's genome, where the first probe comprises a first fluorescent label and a first fluorescence quencher attached to its separate nucleotide residues and the target portion includes the nucleotide residue located at position 69 of Adb797901), and a pair of primers for amplifying a reference portion of the FGF-3 gene, where the reference portion includes the nucleotide residue located at position 69 of Adb797901) and a microarray having at least one oligonucleotide probe that can anneal with a target portion of a mammal's genome, where the target portion includes the nucleotide residue located at position 69 of Adb797901. The method is useful for detecting SNP in FGF-3 gene in a mammal, preferably a human, and is also useful for assessing the relative susceptibility of a mammal to cancer (especially oesophageal cancer), which shows an association with the presence of the C-allele. The cancer is chosen from oesophageal, breast, ovarian, prostate, head and neck cancer. The oesophageal cancer is oesophageal squamous cell carcinoma. The present sequence is a Taqman PCR probe used to genotype individuals for the C SNP in the UTR of FGF-3.

AA	Sequence
22 BP; 3 A; 12 C; 3 G; 4 T; 0 U; 0 Other;	50

Query Match	100.0%	Score 22;	DB 13;	Length 22;
Best local Similarity	100.0%	Pred. No. 13;		
Matches 22; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	CTCCCTGCGCTTCAGCCACATG	22
Db	1	CTCCCTGGCTTCAGCCACATG	22

RESULT 2
ADT97906
ID ADT97906 standard; DNA; 22 BP.

AC	ADT97906;
XX	
DT	27-JAN-2005 (first entry)

Human FGF-3 promoter 5' proximal region T allele probe.

KM Human, wt; fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer
 KW breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
 KM oesophageal squamous cell carcinoma; PCR; probe; SNP;
 KM single nucleotide polymorphism.

OS Homo sapiens.

	Key	Location/Qualifiers
FA		
FH	modified base	1
FT		

2	/	22
FT	mod_base= OTHER	
FT	/note= "modified with Vic (not defined) "	
FT	modified base	

```

FT      /mod_base= OTHER
FT      /note= "G is covalently linked to a TANRA moiety (6-
FT      carbotetramethyl-rhodamine)"

```

XX
PN US2004219582-A1.

PD 04-NOV-2004.

PF 11-MAR-2004; 2004US-00798652.

PR 17-MAR-2003; 2003US-0455689P.

PA (GUOY/) GUO Y.

PI Guo Y;

DR WPI; 2004-794435/78.

PT Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
PS Claim 34, SEQ ID NO 6; 15pp; English.

The invention relates to an isolated nucleic acid molecule comprising the upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3 gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at position 69. Also included are an isolated nucleic acid molecule comprising a sequence complementary to ADT97901, a vector comprising ADT97901 operably linked to a reporter gene, a host cell comprising the vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating a nucleic acid sample from the mammal, and determining whether a cytosine or thymine is present at position 69 of ADT97901), a kit for performing the method (comprising a first oligonucleotide probe which anneals specifically with the target portion of the mammal's genome, where the first probe comprises a first fluorescent label and a first fluorescence quencher attached to its separate nucleotide residues and the target portion includes the nucleotide residue located at position 69 of ADT97901, and a pair of primers for amplifying a reference portion of the FGF-3 gene, where the reference portion includes the nucleotide residue located at position 69 of ADT97901) and a microarray having at least one oligonucleotide probe that can anneal with a target portion of a mammal's genome, where the target portion includes the nucleotide residue located at position 69 of ADT97901. The method is useful for detecting SNP in FGF-3 gene in a mammal, preferably a human, and is also useful for assessing the relative susceptibility of a mammal to cancer (especially oesophageal cancer), which shows an association with the presence of the C-allele. The cancer is chosen from oesophageal, breast, ovarian, prostate, head and neck cancer. The oesophageal cancer is oesophageal squamous cell carcinoma. The present sequence is a Taqman PCR probe used to genotype individuals for the T SNP in the UTR of FGF-3.

Sequence 22 BP; 4 A; 12 C; 2 G; 4 T; 0 U; 0 Other;

Query Match	92.7%	Score 20.4;	DB 13;	Length 22;
Best Local Similarity	95.5%;	Pred. No. 60;		
Matches 21; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
|||
Db 1 CTCCTCACCTTCAGGCCACATG 22

RESULT 3
AEA51146
ID AEA51146 standard; DNA; 20 BP.

AC AEA51146;

DT 11-AUG-2005 (first entry)

DE pyranalR primer, SEQ ID NO: 20 to construct pyrF selection marker system.

AA Expression; protein production; ss; pyrF; PCR; primer.
KW

Unidentified.

XX WO2005052151-A1.
 PN
 XX
 PD 09-JUN-2005.
 XX
 PF 19-NOV-2004; 2004WO-US038884.
 XX
 PR 19-NOV-2003; 2003US-0523420P.
 PR 16-JAN-2004; 2004US-0537147P.
 XX
 PA (DOMC) DOM GLOBAL TECHNOLOGIES INC.
 PI Schneider JC, Chew LC, Badgley AK, Ramseder TM;
 DR WPI, 2005-417995/42.
 XX
 PT New auxotrophic Pseudomonas cell for use in a bacterial expression system
 PT that comprises a nucleic acid construct encoding a polypeptide that
 PT restores prototrophy to the host cell, useful for producing of
 PT recombinant polypeptides.
 XX
 PS Example 1; SEQ ID NO 20; 121pp; English.
 XX
 CC The present invention provides an improved expression system for the
 CC production of recombinant polypeptides utilizing auxotrophic Pseudomonas
 CC cell selectable markers. The invention also provides improved recombinant
 CC protein production in host cells through improved regulation of
 CC expression. The present sequence is the PCR primer used to construct pyrF
 CC selection marker system in a Pseudomonas fluorescens host cell expression
 CC system.
 XX
 SQ Sequence 20 BP; 4 A; 10 C; 2 G; 4 T; 0 U; 0 Other;
 Query Match 70.0%; Score 15.4; DB 14; Length 20;
 Best Local Similarity 94.1%; Pred. No. 7e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 5 CTCGCCTCCAGCACATG 21
 Db 1 CTCGCCTCCAGCACATG 17
 RESULT 4
 ADM79585
 ID ADM79585 standard; DNA; 20 BP.
 AC ADM79585;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human ARTS1 related G446A mutation detection primer SEQ ID NO:14.
 XX
 KW human; ARTS1; ADP-riboylation factor-like tumour suppressor 1;
 KW tumour suppressor; chromosome 13; cytosolic;
 KW abnormal cell proliferation inhibitor; caspase-1 protease modulator;
 KW mutation; detection; PCR; primer; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO2004033659-A2.
 PN
 PD 22-APR-2004.
 XX
 PF 10-OCT-2003; 2003WO-US032270.
 XX
 PR 11-OCT-2002; 2002US-0417842P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 PI Croce CM;
 XX
 DR WPI, 2004-330451/30.

XX Novel ADP-riboylation factor-like tumor suppressor 1 (ARTS1) protein
 PT useful for identifying modulators of Caspase-1 protease protein activity
 PT and for preventing abnormal cell growth in mammalian subjects.
 XX
 PS Example; SEQ ID NO 14; 47pp; English.
 XX
 CC The present invention describes the human ARTS1 (ADP-riboylation factor-
 CC like tumour suppressor 1) protein (I). The human ARTS1 gene is located on
 CC chromosome 13, more specifically to 13q14. Also described: (1) an
 CC isolated nucleic acid molecule (II) comprising a sequence that encodes
 CC (I); (2) an isolated nucleic acid molecule (III) comprising the 3791
 CC nucleotide sequence of SEQ ID NO:2, or its fragment having at least 10
 CC nucleotides; (3) a recombinant expression vector (IV) comprising (III);
 CC (4) a host cell comprising (IV); (5) an oligonucleotide molecule (V)
 CC comprising a nucleotide sequence complementary to a nucleotide sequence
 CC of at least 5 nucleotides of SEQ ID NO:2; and (6) an isolated antibody
 CC (VI) that binds to an epitope on SEQ ID NO:1. Human ARTS1 has cytosolic
 CC activity, and can be used as an inhibitor of proliferation of abnormal
 CC cells. (I) is useful for identifying modulators of caspase-1 protease
 CC protein activity which involves performing a test assay by contacting a
 CC caspase-1 protease protein with a caspase-1 substrate such as FXBP46
 CC protein in the presence of a test compound, determining the level of
 CC processing of the substrate by the protease, and comparing the level to
 CC the level of processing of a caspase-1 substrate by caspase-1 protease
 CC protein in the absence of the test compound. (I) and (II) are useful for
 CC preventing abnormal cell growth in mammalian subjects. (II) is useful for
 CC designing probes and primers that are useful for detecting mutants of
 CC ARTS1 protein. (V) is useful as a probe in oligonucleotide hybridisation
 CC methods for identifying ARTS1. The present sequence represents a PCR
 CC primer for the detection of the G446A mutation in human ARTS, which is
 CC used in an example from the present invention.
 XX
 SQ Sequence 20 BP; 4 A; 8 C; 4 G; 4 T; 0 U; 0 Other;
 Query Match 67.3%; Score 14.8; DB 12; Length 20;
 Best Local Similarity 88.9%; Pred. No. 1.2e+04;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 5 CTCGCCTCCAGCACATG 22
 Db 2 CTCGCCTCCAGCACATG 19
 RESULT 5
 ADY00944/C
 ID ADY00944 standard; DNA; 29 BP.
 AC ADY00944;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE PCR primer 170 used to amplify human GALE region SNP DNA.
 XX
 KW SNP detection; breast tumor; endocrine disease;
 KW gynecology and obstetrics; neoplasia; cytosolic; metastasis;
 KW gene therapy; RNA interference; UDP-galactose 4-epimerase; GALE;
 KW UDP-glucose 4-epimerase; lysophospholipase; HT014; LOC148902; LYPLA2; ss;
 KW PCR; primer.
 XX
 OS Homo sapiens.
 OS
 PN WO2005014846-A2.
 PN
 PD 17-FEB-2005.
 XX
 PF 27-MAY-2004; 2004WO-US016939.
 XX
 PR 24-JUL-2003; 2003US-0490234P.
 PR 25-NOV-2003; 2003US-00723681.
 PR 25-NOV-2003; 2003US-0525239P.
 XX
 PA (SEQU-) SEQUENOM INC.

```

XX  Roch RE, Nelson MR, Braun A, Kammerer SM, Reneland R;
PI  Hoyal-Wrightson CR;
XX
DR  WPI, 2005-163257/17.
XX
PT  Identifying risk of, preventing and/or treating breast cancer by
PT  identifying and/or analyzing polymorphic variations in nucleotide
PT  sequences within the human genome.
XX
PS  Example 15; Page 208; 617pp; English.
XX
CC  The invention relates to a novel method for identifying a subject at risk
CC  of breast cancer comprising detecting the presence or absence of a
CC  polymorphic variation associated with breast cancer. The method of the
CC  invention demonstrates cytostatic activity and may be useful for
CC  identifying a risk of, preventing and/or treating breast cancer and
CC  cancer metastasis. The methods may be utilized for gene therapy or RNA
CC  interference. The current sequence is that of a PCR primer of the
CC  invention which was used to amplify a human HT014/LOC148902/LYP1A2
CC  (lysophospholipase II)/GALB (galactose-4-epimerase, UDP-) DNA containing
CC  a single nucleotide polymorphism (SNP).
XX
SQ  Sequence 29 BP; 8 A; 1 C; 16 G; 4 T; 0 U; 0 Other;
XX
Query Match 66.4%; Score 14.6; DB 14; Length 29;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0.
OY  1 CTCCTCGGCTCCAGCCACAT 21
    ||| ||| ||| ||| |||
DB  28 CTCCTCTCTCTCCCTCCACAT 8
XX
RESULT 6
AAT90350
ID  AAT90350 standard; DNA; 30 BP.
XX
AC  AAT90350;
XX
DT  19-JAN-1998 (first entry)
XX
DE  DNA encoding peptide immunoreactive with HuMAb L94.
XX
KW  Tumour associated antigen; melanoma; cancer; therapy; immunotherapy;
XX  diagnosis; prognosis; vaccine; ss.
XX
OS  Homo sapiens.
XX
PN  WO9714795-A1.
XX
PD  24-APR-1997.
XX
PF  19-OCT-1995; 95WO-US013331.
XX
PR  19-OCT-1995; 95WO-US013331.
XX
PA  (WAYN-) WAYNE CANCER INST JOHN.
XX
PI  Irle RF, Kikumoto Y;
XX
DR  WPI, 1997-245112/22.
XX
P-PSDB; AAW26567.
XX
PT  Antibody against polypeptide having C-terminal amino acid sequence
PT  comprising Ala-Pro, Gly-Pro or Pro-Pro - useful to purify polypeptide
PT  which can be useful in human melanoma treatment.
XX
PS  Example 1; Page 40; 87pp; English.
XX
CC  This nucleotide sequence codes for a peptide (see AAW26567) that is
CC  immunoreactive with human monoclonal antibody WCI L94 that reacts to a
CC  human melanoma cell line. It was obtained by screening melanoma M14 and

```

CC	M12 expression libraries with L99. Isolated clones (see AAT90347-51)
CC	encode peptides (AAM26566-73) that share the C-terminal sequence Ala-Pro.
CC	These peptides can be used in an antigen composition to elicit an immune
CC	response, preferably a cytotoxic T lymphocyte response, specifically as a
CC	polyvalent tumour cell vaccine in conjunction with a cancer whole cell
CC	vaccine therapy for human melanoma. The peptides can also be used to
CC	enhance an immune response, treat and diagnose cancer
XX	
SQ	Sequence 30 BP; 5 A; 12 C; 8 G; 5 T; 0 U; 0 Other;
Query Match	64.5%; Score 14.2; DB 2; Length 30;
Best Local Similarity	84.2%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0.	
Oy	4 CCTCGCCTTCAGCCACATG 22 Db 11 CCGGGCCTCCAGCTCATG 29
RESULT 7	
ID	ADB67764/C
AD	ADB67764 standard; DNA; 25 BP.
AC	ADB67764;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	HIV-1 gp120 optimised sequencing primer g120c-SF4.
XX	
KM	HIV, Tat; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV,
KW	primer; ss.
XX	
OS	Synthetic.
OS	Human immunodeficiency virus 1.
XX	
PN	WO2003011334-A1.
PD	
PD	13-FEB-2003.
XX	
PF	26-JUL-2002; 2002WO-EP008343.
XX	
PR	27-JUL-2001; 2001GB-00018367.
PA	(GLAX) GLAXOSMITHKLINE BIOLOGICALS SA.
PA	(GLAX) GLAXO GROUP LTD.
XX	
PI	Ertl PF, Tite JP, Van Wely CA, Voss G;
DR	WPI; 2003-239474/23.
PS	
Example 15; Fig 21; 108bp; English.	
The present invention describes the use of an HIV Tat protein or	
polynucleotide, an HIV Nef protein or polynucleotide, or an HIV Tat	
protein or polynucleotide linked to an HIV Nef protein or polynucleotide,	
and an HIV sp120 protein or polynucleotide in manufacturing a vaccine for	
a prime-boost delivery for the prophylactic or therapeutic immunisation	
of humans against HIV. The protein or polynucleotide is delivered via a	
bombardment approach. Also described: (1) a recombinant DNA molecule	
comprising a Nef and/or Tat and/or gp120 gene in a vector in which the	
gene of interest is inserted 3' to an enhanced HCMV IRL promoter; (2)	
particles, preferably gold particles, coated with recombinant DNA	
comprising at least two different vaccine compositions including: (a) a	
composition comprising particles coated with DNA encoding gp120 and nef	
and/or tat or nef/tat, and (b) a composition comprising gp120 and nef	
and/or tat or nef-tat DNA or proteins, where the DNA or proteins are not	
coated onto the particles. The HIV proteins or polynucleotides are used	
to produce vaccines having anti-HIV activity. The recombinant DNA	

CC molecule can be used in manufacturing an HIV vaccine for the prophylactic
CC or therapeutic immunisation of humans. The present sequence represents an
CC oligonucleotide primer for sequencing optimised gp120, which is used in
CC an example from the present invention.

SO Sequence 25 BP; 6 A; 4 C; 12 G; 3 T; 0 U; 0 Other;

Query Match 63.6%; Score 14; DB 10; Length 25;
Best Local Similarity 77.3%; Pred. No. 2.7e+04;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CTCGCTCGCTCCAGCCACATG 22
DB 22 CTTTCCACACTCTCTCCACATG 1

RESULT 8

AEAS1123
ID AEAS1123 standard; DNA; 20 BP.

AC AEAS1123;

XX 11-AUG-2005 (first entry)

DE Human class II bHLH protein SCX gene specific sense primer.

XX Cell therapy; stem cell; trauma; tranquilizer or vulnerary; injury;

KW periodontal disease; antiinflammatory; mouth disease; scleraxis gene;

XX primer; ss.

OS Homo sapiens.

PN WO2005052140-A2.

XX 09-JUN-2005.

XX 22-NOV-2004; 2004WO-US039248.

XX 20-NOV-2003; 2003US-052602P.

XX (USSH) US DEPT OF HEALTH.

PI Shi S, Seo B, Miura M;

XX WPI; 2005-405376/41.

PT Novel isolated periodontal ligament multipotent stem cell, useful for
PT generating periodontal tissue or cells in a subject to reduce or
PT ameliorate trauma such as periodontal disease or physical injury due to
PT dental procedure.

XX Example 1; SEQ ID NO 7; 62pp; English.

CC The present invention relates to postnatal periodontal ligament stem
CC cells (PDLSCs). The invention is useful for periodontal tissue
CC regeneration, periodontal disease treatment and differentiation of the
CC cells and methods of tissue cryopreservation. The invention is also
CC useful in cell therapy and in medical therapy e.g. for treating trauma.
CC The present sequence is the human class II bHLH protein scleraxis (SCX)
CC gene specific primer. This sequence is used in the isolation and
CC characterization of PDLSCs.

SO Sequence 20 BP; 3 A; 9 C; 3 G; 5 T; 0 U; 0 Other;

Query Match 62.7%; Score 13.8; DB 14; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 CTCGCTCCAGCCACAT 21
DB 1 CTGGCTCCAGCTACAT 17

RESULT 9

ABQ00949/C
ID ABQ00949 standard; DNA; 24 BP.

XX ABQ00949;

XX 11-JUN-2002 (first entry)

DE Oligonucleotide adapter/capture probe 940.

XX Oligonucleotide array; adapter sequence; probe; ss.

XX Synthetic.

PN WO200216649-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026519.

XX 25-AUG-2000; 2000US-0227948P.

XX 29-AUG-2000; 2000US-0228854P.

XX (ILLU-) ILLUMINA INC.

PI Gunderson K;

XX WPI; 2002-292068/33.

PT Array comprising adapter sequences useful for immobilizing or detecting a
PT target nucleic acid sequence, has different addresses comprising
PT different specific capture probes.

PS Claim 1; Page 66; 261pp; English.

CC The invention relates to an oligonucleotide array (I) comprising at least
CC 25 different addresses (adapter sequences) with each comprising a
CC different capture probe selected from a group consisting of the sequences
CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
CC nucleic acid sequence by attaching a adapter nucleic acid (ABQ0010-
CC ABQ13409) to a target nucleic acid to form a modified target nucleic acid
CC and contacting the modified target nucleic acid with (I). The steps of
CC above method is useful for detecting a target nucleic acid, which further
CC comprises detecting the presence of the modified target nucleic acid

SO Sequence 24 BP; 5 A; 4 C; 8 G; 7 T; 0 U; 0 Other;

Query Match 62.7%; Score 13.8; DB 6; Length 24;
Best Local Similarity 88.2%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 CTCGCTCCAGCCACAT 21
DB 21 CTTCCGCTCCAGACACAT 5

RESULT 10

ABQ05868
ID ABQ05868 standard; DNA; 24 BP.

XX ABQ05868;

XX 11-JUN-2002 (first entry)

DE Oligonucleotide adapter/capture probe 5859.

XX Oligonucleotide array; adapter sequence; probe; ss.

XX Synthetic.

PN WO200216649-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026519.
XX
XX 25-AUG-2000; 2000US-0227948P.
XX 29-AUG-2000; 2000US-0228854P.
XX
XX (ILLU-) ILLUMINA INC.
XX
XX Gunderson K;
XX
XX WPI; 2002-292068/33.
XX
XX Array comprising adapter sequences useful for immobilizing or detecting a
XX target nucleic acid sequence, has different addresses comprising
XX different specific capture probes.
XX
XX Claim 1, Page 161; 261pp; English.
XX
XX The invention relates to an oligonucleotide array (I) comprising at least
XX 25 different addresses (adapter sequences) with each comprising a
XX different capture probe selected from a group consisting of the sequences
XX given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
XX nucleic acid sequence by attaching a adapter nucleic acid (ABQ00010-
XX ABQ13409) to a target nucleic acid to form a modified target nucleic acid
XX and contacting the modified target nucleic acid with (I). The steps of
XX above method is useful for detecting a target nucleic acid, which further
XX comprises detecting the presence of the modified target nucleic acid
XX
XX Sequence 24 BP; 7 A; 8 C; 4 G; 5 T; 0 U; 0 Other;
XX
XX Query Match 62.7%; Score 13.8; DB 6; Length 24;
XX Best Local Similarity 88.2%; Pred. No. 3.2e+04;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 5 CTCGCCTCCAGCACAT 21
XX 4 CTCGCCTCCAGCACAT 20
XX
XX DB
XX
XX RESULT 11
XX ABQ05827/c
XX ID ABQ05827 standard; DNA; 24 BP.
XX
XX AC ABQ05827;
XX
XX 11-JUN-2002 (first entry)
XX
XX Oligonucleotide adapter/capture probe 5818.
XX
XX Oligonucleotide array; adapter sequence; probe; ss.
XX
XX Synthetic.
XX
XX WO200216649-A2.
XX
XX 28-FEB-2002.
XX
XX 27-AUG-2001; 2001WO-US026519.
XX
XX 25-AUG-2000; 2000US-0227948P.
XX 29-AUG-2000; 2000US-0228854P.
XX
XX (ILLU-) ILLUMINA INC.
XX
XX PI Gunderson K;
XX
XX WPI; 2002-292068/33.
XX
XX Array comprising adapter sequences useful for immobilizing or detecting a
XX target nucleic acid sequence, has different addresses comprising
XX different specific capture probes.
XX
XX Claim 1, Page 161; 261pp; English.

XX The invention relates to an oligonucleotide array (I) comprising at least
XX 25 different addresses (adapter sequences) with each comprising a
XX different capture probe selected from a group consisting of the sequences
XX given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
XX nucleic acid sequence by attaching a adapter nucleic acid (ABQ00010-
XX ABQ13409) to a target nucleic acid to form a modified target nucleic acid
XX and contacting the modified target nucleic acid with (I). The steps of
XX above method is useful for detecting a target nucleic acid, which further
XX comprises detecting the presence of the modified target nucleic acid
XX
XX Sequence 24 BP; 5 A; 4 C; 8 G; 7 T; 0 U; 0 Other;
XX
XX Query Match 62.7%; Score 13.8; DB 6; Length 24;
XX Best Local Similarity 88.2%; Pred. No. 3.2e+04;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 5 CTCGCCTCCAGCACAT 21
XX 21 CTCGCCTCCAGCACAT 5
XX
XX DB
XX
XX RESULT 12
XX AAS07411
XX ID AAS07411 standard; DNA; 20 BP.
XX
XX AC AAS07411;
XX
XX 12-SEP-2001 (first entry)
XX
XX PCR primer K21L19L used to map the REVOLUTA gene.
XX
XX DE Revoluta; Rev; corn; barley; rice; tomato; PCR primer; apical meristem;
XX leaf; floral organ; stem; transgenic plant; crop yield; cereal; fruit;
XX pharmaceutical; industrial; ss.
XX
XX Arabidopsis thaliana.
XX
XX OS WO20013944-A1.
XX
XX PN 17-MAY-2001.
XX
XX PD 10-NOV-2000; 2000WO-US030794.
XX
XX PR 10-NOV-1999; 99US-0164587P.
XX
XX PA (SLAD/) SLADE A.
XX PA (MADI/) MADISEN L.
XX PA (COMA/) COMAI L.
XX
XX PI Slade A, Madisen L, Comai L;
XX
XX WPI; 2001-328861/34.
XX
XX Isolated DNA molecule comprising a sequence that encodes a REVOLUTA
XX protein, useful for producing transgenic plants with modulated cell
XX division.
XX
XX Example 1; Page 48; 149pp; English.
XX
XX AAS07401-AAS07571 represent REVOLUTA (REV) coding sequences and PCR
XX primers of the invention. The REV nucleic acid sequences were isolated
XX from plants such as Arabidopsis thaliana, tomato, corn, barley and rice.
XX The REV gene is required to promote the growth of apical meristems, but
XX has an opposite effect on meristems of leaves, floral organs and stems,
XX such that it acts to limit cell division reducing the rate of plant
XX growth and final size of the tissue. Therefore, loss of functional REV
XX leads to increases in the size of floral organs, leaf and stem tissue.
XX CC DNA encoding the REV protein is useful for modulating plant cell
XX division. The mutant REV DNA is also useful for producing transgenic
XX plants with modulated cell division. These transgenic plants can be used
XX to increase crop yield in cereals and fruits, and as a potential source
XX of pharmaceuticals and industrial products

XX Sequence 20 BP; 4 A; 10 C; 1 G; 5 T; 0 U; 0 Other;
SQ

Query Match 61.8%; Score 13.6; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 1 CTCCTCTTTCACAGACACA 20

RESULT 13
AA57861
ID AA57861 standard; DNA; 22 BP.
XX
AC AA57861;
XX
DT 15-JUL-1999 (first entry)
XX
DE PCR primer used in construction of yeast artificial chromosome.
XX
KM YAC; yeast artificial chromosome; PCR primer; sexual dysfunction;
KM reporter gene; transgenic mammal; therapy; circadian function;
KM sleep disorder; eating disorder; premenstrual syndrome; birth defect;
KM autoimmune disorder; ss.
XX
OS Synthetic.
XX
PN GB2331752-A.
XX
PD 02-JUN-1999.
XX
PF 27-NOV-1998; 98GB-00026126.
XX
PR 28-NOV-1997; 97GB-00025311.
PR 28-NOV-1997; 97GB-00025313.
PR 20-MAR-1998; 98GB-00006072.
PR 05-NOV-1998; 98GB-00024275.
XX
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Shen S, Schedl A, Harmar AJ;
XX
DR WPI; 1999-290603/25.
XX
PT New reporter gene labeled YAC vectors and transgenic mammals used for
PT screening potential active agents.
XX
PS Dielosure; Page 56; 98pp; English.
XX
CC This sequence represents a PCR primer used in the construction of a yeast
CC artificial chromosome of the invention. The yeast artificial chromosome
CC (YAC) vectors contain a reporter gene and transgenic mammals produced
CC using them may be used to screen for an agent affecting nucleotide
CC expression and gives easier monitoring of in vivo expression. The vector
CC is used in the production of transgenic mammals for testing potential
CC pharmaceutical or veterinary agents. pYAM4 is used to amplify YAC. The
CC assay may be used to screen for agents useful in treatment of disturbance
CC of circadian function, sleep disorders, eating disorders, premenstrual
CC syndrome, autoimmune disorders, birth defects in women and/or sexual
CC dysfunction. The agents thus detected may be used for treatment of
CC disorders related to the expression pattern of a nucleotide such as those
CC above. The vectors have more concentrated YAC DNA, which allows better
CC and more reliable gene transfer. The presence of a reporter gene allows
CC easy monitoring of in vivo expression and the vectors allow for gene
CC overexpression (3-5 fold) and easy site determination. The pYAM4
CC amplification vector does not contain the thymidine kinase gene, which
CC causes male infertility in transgenic mice
XX
SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
Query Match 61.8%; Score 13.6; DB 2; Length 22;

Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TCCTCGCCTCCAGCCACAT 21
Db 1 TTCATCACCTCCATCCACAT 20

RESULT 14
AAC85242
ID AAC85242 standard; DNA; 22 BP.
XX
AC AAC85242;
XX
DT 22-MAR-2001 (first entry)
XX
DE Reverse primer D 26376 for determining size of YAC insert.
XX
KM Internal ribosomal entry site; IRES; yeast artificial chromosome; YAC;
KM vector; centromere; telomere; origin of replication; transgenic;
KM circadian function; sleep disorder; eating disorder;
KM premenstrual syndrome; autoimmune disease; birth defect;
KM sexual dysfunction; serotonin transporter; VIP2 receptor; SERT; VIPR2;
KM polymerase chain reaction; PCR; primer; amplify; YAC 35D8/D6;
KM YAC HSC7B526/V12; ss.
XX
OS Synthetic.
XX
PN GB2350613-A.
XX
PD 06-DEC-2000.
XX
PF 17-AUG-2000; 2000GB-00020335.
XX
PR 28-NOV-1997; 97GB-00025311.
PR 28-NOV-1997; 97GB-00025313.
PR 20-MAR-1998; 98GB-00006072.
PR 05-NOV-1998; 98GB-00024275*
PR 27-NOV-1998; 98GB-00026126.
XX
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Shen S, Schedl A, Harmar AJ;
XX
DR WPI; 2001-034098/05.
XX
PT Transgenic organism for identifying potential therapeutic agents able to
PT modulate gene expression, comprises a yeast artificial chromosome vector.
XX
PS Example; Page 54; 93pp; English.
XX
CC The sequences given in AAC85227-50 are primers which were used to
CC determine the size of the integrated YAC 35D8/D6 and YAC HSC7B526/V12
CC constructs in a transgenic founder animal. The constructs were prepared
CC from novel yeast artificial chromosome (YAC) vectors each of which
CC comprises a centromere, two telomeres, at least one origin of
CC replication, an internal ribosomal entry site (IRES), and a selection
CC gene that is specifically removable from the vector. The resulting YAC
CC are used to produce transgenic organisms for use in screening for agents
CC that can affect the expression pattern of a nucleotide sequence of
CC interest (NOI) or the activity of its expression product. The identified
CC agents are potentially useful as pharmaceutical and veterinary agents for
CC treating disorders of circadian function; sleep or eating disorders;
CC premenstrual syndrome; autoimmune diseases; birth defects in women and/or
CC sexual dysfunction, and also as lead compounds for developing agents with
CC other activities. YAC can also be used for expression, regulation and/or
CC functional studies on NOI, in combination with other NOI, compounds or
CC compositions. The new vectors provide high YAC copy numbers and allow
CC easy monitoring (in vivo) of the expression pattern of NOI;
CC (over)expression of NOI and a reporter gene; and determination of the
CC sites where NOI is expressed. Incorporation of IRES allows expression of
CC at least two nucleic acid sequences (e.g. NOI plus a reporter of more
CC than one NOI)

XX SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
Query Match 61.8%; Score 13.6; DB 5; Length 22;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TCCCTCGCCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 TTCATCACCCTCCATCCACAT 20
| | | | | | | | | | | | | | | | | | | | | |
RESULT 15
ACI88529/c
ID ACI88529 standard; DNA; 25 BP.
XX ACI88529;
XX
XX 14-OCT-2003 (first entry)
XX
XX Human microarray DNA oligonucleotide SEQ ID NO 88520.
XX
XX EST; ss; probe; expressed sequence tag; microarray; gene expression;
XX genetic variation; biallelic marker; polymorphism; human;
XX cross-species comparison.
XX
XX Homo sapiens.
XX
XX US2003104410-A1.
XX
XX 05-JUN-2003.
XX
XX 15-MAR-2002; 2002US-00098263.
XX
XX 16-MAR-2001; 2001US-0276759P.
XX
XX (AFY-) AFFYMETRIX INC.
XX
XX Miltmann MP;
XX
XX WPI; 2003-567953/53.
XX
XX New array of nucleic acid probes, useful for in situ hybridization, in
XX Southern, Northern or dot-blot hybridization to identify or detect the
XX sequence or specific mutations of any gene.
XX
XX Claim 1; SEQ ID NO 88520; 9pp; English.
XX
XX The invention discloses a microarray comprising a plurality of nucleic
XX acid probes including one of 2,018,500 fully defined sequences, or its
XX perfect match, perfect mismatch, antisense match or antisense mismatch.
XX Also disclosed is a method of gene expression analysis. The array is used
XX in monitoring gene expression levels by hybridisation to a DNA library,
XX in analysis of genetic variation or in hybridisation of tag-labelled
XX compounds. The nucleic acid probes are specifically designed for analysis
XX of at least one target sequence. The method of analysis comprises
XX hybridising at least one or more nucleic acids to at least two or more
XX nucleic acid probes and detecting the hybridisation. The nucleic acid
XX probes are attached to a solid support. The analysis comprises monitoring
XX gene expression levels, identifying biallelic markers or polymorphisms,
XX or family members of a gene and a cross-species comparison. Each of the
XX nucleic acids further comprises a tag sequence. The array of nucleic acid
XX probes is useful in situ hybridisation, in Southern, Northern or dot-
XX blot hybridisation to identify or detect the sequence or specific
XX mutations of any gene, in mapping the 5' termini of mRNA molecules by
XX primer extensions or in screening cDNA or genomic libraries or subclones
XX for additional subclones containing segments of DNA that have been
XX isolated and previously sequenced. The sequence presented is one of the
XX nucleic acid probes incorporated in the microarray. Note: The sequence
XX data for this patent can also be obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html
XX
XX Sequence 25 BP; 9 A; 3 C; 10 G; 3 T; 0 U; 0 Other;
XX SQ

Query Match 61.8%; Score 13.6; DB 9; Length 25;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TCCCTCGCCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 23 TCCCTCGCTCCACCTACTT 4
| | | | | | | | | | | | | | | | | | | | | |
Search completed: March 25, 2006, 19:02:19
Job time : 374.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 13:43:43 ; Search time 3041.5 Seconds
(without alignments)
338.424 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 ctccctcgccgcacgacacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 52094

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hnc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsal:*
10: gb_gsal2:*
11: gb_gsal3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14.2	64.5	25	10	CG730897 1119130F0
C 2	12.4	56.4	29	10	CG676355 PR10118B
C 3	12	54.5	24	9	AZ762096 1M055618
C 4	12	54.5	25	1	A1416870 sal1b03.x
C 5	12	54.5	25	8	D21049
C 6	12	54.5	29	9	AZ332711 1M0061N4
C 7	12	54.5	29	9	AZ788256 2M0035A17
C 8	12	54.5	29	10	C2477475
C 9	11.6	52.7	22	9	AZ980090 2M0258F17
C 10	11.4	51.8	24	9	AZ655652 1M0530D13
C 11	11.4	51.8	28	1	AA864650 ch37b09.8
C 12	11.4	51.8	28	1	A1158446
C 13	11.2	50.9	16	1	AM247040 2822414.5
C 14	11.2	50.9	22	9	AZ863707 2M0171N19
C 15	11.2	50.9	28	1	A1279523 q151d12.x
C 16	11.2	50.9	29	9	AZ498825 1M0336J14
C 17	11	50.0	22	2	BG926061 HNC23-1-E
C 18	11	50.0	22	9	AZ778032 2M0013A11
C 19	11	50.0	23	9	AZ785027 2M0028H03
C 20	11	50.0	24	9	AZ404465 1M0172P09
C 21	11	50.0	25	1	A1424019 t51e01.x
C 22	11	50.0	27	9	AZ439789 1M0230004

23	11	50.0	27	9	BH851753
C 24	11	50.0	28	9	AZ937148
C 25	11	50.0	29	9	AZ666737
C 26	11	50.0	29	9	AZ780387
C 27	11	50.0	29	11	CT024781
C 28	11	50.0	30	2	BF099623
C 29	11	50.0	30	2	BE367782
C 30	11	50.0	30	8	DN955616
C 31	11	50.0	30	9	AZ634665
C 32	11	50.0	30	9	AZ864869
C 33	10.8	49.1	19	9	AZ485264
C 34	10.8	49.1	21	9	AZ348213
C 35	10.8	49.1	23	9	AZ627985
C 36	10.8	49.1	27	9	AZ797359
C 37	10.8	49.1	28	5	C00171
C 38	10.8	49.1	28	9	AZ343801
C 39	10.8	49.1	29	9	AZ608734
C 40	10.8	49.1	30	8	CV999848
C 41	10.8	49.1	30	9	AZ456295
C 42	10.8	49.1	30	10	C2194668
C 43	10.6	48.2	19	9	AZ381798
C 44	10.6	48.2	19	9	AZ782026
C 45	10.6	48.2	20	7	CF932153

ALIGNMENTS

RESULT 1
LOCUS CG730897/c 25 bp DNA linear GSS 20-OCT-2003
DEFINITION 1119130F05.2EL x1 1119 - Rescemu Grid AA Zea mays genomic, genomic survey sequence.

ACCESSION CG730897
VERSION CG730897.1 GI:37773389
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 25)

AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered Rescemu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot, V.
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu

Reverse complemented post-ligation sequence from source sequence.
Plate: 1119130 row: F column: 05
Class: transposon-tagged.
Location/Qualifiers
1..25
/organism="Zea mays"
/mol_type="genomic DNA"
/culturvar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/feature_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="1119 - Rescemu Grid AA"
/note="Organ: leaf; Vector: Rescemu Mu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; Rescemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescemu, go to the web site 'www.zmbd.iastate.edu' and follow the links for

FEATURES

source

ACCESSION	A1416870
VERSION	A1416870.1
KEYWORDS	GI:4260374
SOURCE	EST.
ORGANISM	Glycine max (soybean)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycyde. 1 (bases 1 to 25) Shenmaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coveyell, V., Khanna, A., Bola, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepietz, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, J., Cardenas, M., McCam, R., Waterston, R. and Wilson, R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu When it has been determined, an EST from the other end of this clone is listed in the "Other ESTs on clone" field. GENOME SYSTEMS CLONE ID: Gm-cl005-102 Trace considered overall poor quality Possible reversed clone: similarity on wrong strand This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 5706 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Seq primer: 17 ET from Amesham High quality sequence atop: 1 POLYA-No.

```

FEATURES
Source
Location/Qualifiers
1. .25
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1005-102"
/lab_host="XLI0-Gold"
/clone_1lb="Gm-c1005"
/notes="Vector: pBluescript II XR; Site.1: EcoRI; Site.2:
XhoI; This library was constructed by Dr. Randy C.
Shoemaker and Dr. John Expelling, USDA-ARS Agronomy
Department, G401 Agronomy Hall, Iowa State University,
Ames, IA 50011-1010 Phone: 515-294-6233 FAX: 515-294-2299
email: rcs8shoe@iastate.edu"

```

Query Match	54.5%	Score 13	DB 1	Length 25
Best Local Similarity	75.0%	Pred. No. 1.7e+06		
Matches 15	Conservative 0	Mismatches 5	Indels 0	Gaps 0
Qy	1	CTCCCTGCGCTTCACGCCACA	20	
Db	6	CCCCCCCCCTTCACACCCCCCA	25	

RESULT	5
D21049	
LOCUS	25 bp mRNA linear EST 30-JUL-1996
DEFINITION	HUNGSD2034 Human promyelocyte Homo sapiens cDNA clone mp0338 3'
ACCESSION	D21049
VERSION	D21049.1 GI:504869
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
1 (bases 1 to 25)	Okubo, K., Fukushima, A., Yoshii, J., Niiyama, T., Kojima, Y., Yoshinari, H., Arimoto, J., and Matsubara, K.	Gene expression of human promyelocytic cell line HL60 before and after induction of differentiation. A new application of 3'directed cDNA sequencing	Unpublished (1993)	Contact: Okubo, K., Fukushima, A., Yoshii, J., Niiyama, T., Kojima, Y., Yoshinari, H., Arimoto, J., and Matsubara, K. Institute for Molecular and Cellular Biology Osaka University 3-1 Yamada-oka, Suita, Osaka 565, Japan.	location/Qualifiers 1..25 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="mp0338" /note="Femalce, adult, cell_line = HU60, cell_type = promyelocyte."

Query Match	54.5%;	Score 12;	DB 8;	Length 25;
Best Local Similarity	75.0%;	Pred. No. 1.7e+06;		
Matches	15;	Conservative	0;	Mismatches 5;
				Indels 0;
				Gaps 0;
Oy	1	CTCGCTGCGCTCCAGCCACA	20	
Db	4	CGCCCCACCTCCACCTCTCA	23	

RESULT 6	AZ332711/c
LOCUS	AZ332711 29 bp DNA Linear GSS 29-SEP-2000
DEFINITION	1M0061N1F Mouse 10kb plasmid UOCCIM library Mus musculus genomic
ACCESSION	AZ332711
VERSION	AZ332711
KEYWORDS	AZ332711.1 GI:10396624
SOURCE	GSS.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus. 1 (bases 1 to 29) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,B., Petersen,T., Rellay,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright D.Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contact: Robert B. Weiss
JOURNAL	
COMMENT	

```

FEATURES
source
    1. .29
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="U96C1M0061N14"
        class=plasmid ends
        High quality sequence stop: 29.
        Location/Qualifiers

```

/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCGIM library"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The sheared DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 54.5%; Score 12; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CTCGAGCCACAT 21
| | | | | | | | | |
Db 28 CTCGAGCCACAT 17

RESULT 7

A2788256/c

LOCUS 29 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0035A17F Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCG2M0035A17 F, genomic survey sequence.

A2788256

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sclerogomphi; Muridae; Murinae; Mus.

1 (bases 1 to 29)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenan, E., Pedersen, T.,

Reilly, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112 USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0035 row: A column: 17

Seq primer: CGTGTAAACGACGCGCAT

Class: plasmid ends

High quality sequence sloop: 29.

Location/Qualifiers

1..29

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UGCG2M0035A17"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCGIM library"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The sheared DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 54.5%; Score 12; DB 9; Length 29;
Best Local Similarity 75.0%; Pred. No. 1.7e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CCTCGCCTCCAGCCACATG 22
| | | | | | | | | |
Db 29 CACCCCTCCGCGCACAG 10

RESULT 8

C2477475

LOCUS 29 bp DNA linear GSS 29-APR-2005
DEFINITION d10866-3prine Exelixis P element XP insertions Drosophila
melanogaster genomic Sequence recovered from 3' end of P element,
genomic survey sequence.

C2477475

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 29)

Thibault, S.T., Singer, M.A., Miyazaki, W.Y., Mlasek, B., Dempe, N.A.,

Singh, C.M., Buchholz, R., Demsky, M., Fawcett, R., Francis-Lang, H.L.,

Ryer, J., Cheung, L.M., Chong, A., Erickson, C., Fisher, W.W.,

Greer, K., Hartouni, S.R., Howie, E., Jakula, L., Joo, D., Killpack, K.,

Laufer, A., Mazzotta, J., Smith, R.D., Stevens, L.M., Stuber, C.,

Tan, L.R., Ventura, R., Woo, A., Zakrzewski, I., Zhao, L., Chen, F.,

Swammer, C., Kopiczynski, C., Duyk, G., Winberg, M.L. and Margolis, J.

A complementary transposon tool kit for Drosophila melanogaster

using P and piggyBac

Nat. Genet. 36 (3), 283-287 (2004)

14981521

Contact: Roger A Hoskins

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory

Mailstop 64-121, One Cyclotron Road, Berkeley, CA 94720, USA

Tel: 510 486 4015

Fax: 510 486 6798

Email: Rhooskins@lbl.gov

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P

element.

The P element insertion position is 1 in the 29 bases. This

insertion position refers to the first base of the 8 base target

recognition sequence.

FEATURES	source	location/Qualifiers	Class: transposon insertion site.
		1..29	
		/organism="Drosophila melanogaster"	
		/mol_type="genomic DNA"	
		/strain="isogenic w- strain"	
		/db_xref="taxon:7227"	
		/clone_idb="Exelixis P element XP insertions"	
		/note="Vector: P element XP (GenBank accession number AY515149); An isogenic w- Drosophila melanogaster strain was mutagenized by remobilization of transposable elements. For the P element XP, we selected an easily mobilized ammunition element among inserts hopped onto the Blominey balancer. New insertions were collected in vias from dysgenic females using the standard chromosomal source of transposase, delta2-3. All lines were mapped to a chromosome by standard genetic methods, examined for homozygous viability and used for recovery of flanking genomic sequence by inverse PCR."	
ORIGIN			
Query Match		54.5%; Score 12; DB 10; Length 29;	
Best Local Similarity		75.0%; Pred. NO. 1.7e+06;	
Matches	15; Conservative	0; Mismatches	5; Indels
			0; Gaps
0y			
	1 CTCCTGCGCTTCAGCCACA	20	
	6 CTCGAGTTCCTCAGCCACA	25	
RESULT 9			
AZ980990/c			
LOCUS	AZ980990	22 bp	DNA
DEFINITION	2M0258F17F Mouse 10kb plasmid UUGC2M 1library Mus musculus genomic		
ACCESSION	clone UUGC2M0258F17 F, genomic survey sequence.		
VERSION	AZ980990		
KEYWORDS	AZ980990.1 GI:13852217		
SOURCE	GSS.		
ORGANISM	Mus musculus (house mouse)		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciuromorphi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 22)		
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,		
	Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T.,		
	Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von		
	Niederhausen,A. and Wright,D.,Weiss,R.		
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb		
	plasmid inserts		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Robert B. Weiss		
	University of Utah Genome Center		
	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT		
	84112, USA		
	Tel: 801 585 5606		
	Fax: 801 585 7177		
	Email: ddunn@genetics.utah.edu		
	Insert Length: 10000 Std Error: 0.00		
	Plate: 0258 row: F column: 17		
	Seq primer: CGTGTAAACGACGCGCAGT		
	Class: plasmid ends		
FEATURES	High quality sequence stop: 22.		
source	location/Qualifiers		
	1..22		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="UUGC2M0258F17"		
	/sex="Female"		
	/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"		

```

ORIGIN
Query Match      52.7%  Score 11.6;  DB 9;  Length 22;
Best Local Similarity 77.8%;  Pred. No. 2.4e+06;
Matches 14;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;

Db      22  CCCCACCTCCACACACA 5

QY      3  CCTTCGCTCCAGCCACA 20
      ||| ||| ||| ||| |||
      ||| ||| ||| ||| |||

RESULT 10
LOCUS      AZ655653/c
DEFINITION 1M0530D13R Mouse 10kb plasmid UUGCIM1 library Mus musculus genomic
ACCESSION  AZ655652
VERSION     AZ655652
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus (house mouse)
REFERENCE   Bukaryote, Metazoa; Chordata, Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
            1 (bases 1 to 24)
            Dunn, D., Aoyagi, A., Barber, M., Beacom, T., Duval, B., Hamil, C.,
            Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
            Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
            Niederhausern, A. and Wright, D., Weis, R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
TITLE       Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel.: 801 585 5606
            Fax: 801 585 7177
            Email: edunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0530  row: D  column: 13
            Seq primer: CACACAGCAAAACAGCATATGACC
            Class: plasmid ends
            High quality sequence stop: 24.
            Location/Qualifiers
                1..24
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGCIM0530D13"
                /sex="Male"

```


RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN

Query Match 51.8%; Score 11.4; DB 1; Length 28;
Best Local Similarity 92.3%; Pred. No. 3e+06;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 7 CGCCTCCAGCCAC 19
|||||
Db 10 CGCCTACAGCCAC 22

RESULT 13

AM247040 16 bp mRNA linear EST 07-JAN-2000
LOCUS AM247040/c
DEFINITION 2822414.Sprine NIH_MGC_7 Homo sapiens CDNA clone IMAGE:2822414 5',
mRNA sequence.

ACCESSION AM247040
VERSION AM247040
KEYWORDS GI:6590033
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS 1 (bases 1 to 16)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Other ESTs: 2822414.3prine
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DP CDNA Library Preparation: Ling
Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing
project clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
trimming: cross match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu Low Quality Sequence: 15
contiguous PHRED high quality bases following vector sequence. Very
Low Quality Sequence: Trace file contained 16 contiguous distinct
peaks following vector sequence.
Plate: LCM9 row: F column: 15
High quality sequence stop: 15.

FEATURES

Source

1. 16
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2822414"
/issue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 7"
/note="Organ: lung; Vector: pOTB7, Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 50.9%; Score 11.2; DB 1; Length 16;
Best Local Similarity 81.2%; Pred. No. 3.6e+06;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 TCCTCGCCTCCAGCC 17
|||||
Db 16 TCCTTCGCTCTGTC 1

RESULT 14

AZ863707 22 bp DNA linear GSS 21-FEB-2001
LOCUS AZ863707/c
DEFINITION 2M0171N19R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCG2M0171N19 R, genomic survey sequence.

ACCESSION AZ863707
VERSION AZ863707.1 GI:13062279
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

REFERENCE
AUTHORS Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 22)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausen,A. and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0171 row: N column: 19
Seq primer: CACACAGGAACAGCTATGACC
Clase: plasmid ends
High quality sequence stop: 22.

FEATURES

Source

1. 22
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0171N19"
/sex="Male"
/lab_host="B. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UGCGIM library"
/note="Vector: pMD22nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD22 (g14732114[g14732072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 50.9%; Score 11.2; DB 9; Length 22;
Best Local Similarity 81.2%; Pred. No. 3.6e+06;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 CCTGCGCTCCAGCCA 18
 |||||
 19 CCATGCGCGCCACCA 4

RESULT 15
 A1279523 28 bp mRNA linear EST 29-JAN-1999
 LOCUS A1279523
 DEFINITION Q151d12.x1 Soares NHHMPu S1 Homo sapiens cDNA clone IMAGE:1875863
 3, similar to SW:EXTN_TOBAC P13983 EXTENSIN PRECURSOR ; contains
 MER22.b1 MSRI repetitive element ;, mRNA sequence.

ACCESSION A1279523
 VERSION A1279523
 KEYWORDS EST, GI:3917757
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE 1 (bases 1 to 28)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Insert Length: 588 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES
 source location/Qualifiers

1..28
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1875863"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares NHHMPu S1"
 /note="Organ: mixed (see below); Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte ZNHM, pregnant uterus
 NHHMPu, and fetal heart NHH19W) were mixed, and as circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-345479, and 484488-489479."

ORIGIN

Query Match 50.9%; Score 11.2; DB 1; Length 28;
 Best Local Similarity 81.2%; Pred. No. 3.6e+06;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 4 CCTGCGCTCCAGCCAC 19
 |||||
 12 CCTGCGCGCGCCAC 27

Search completed: March 25, 2006, 15:28:19
 Job time : 3043.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 12:58:33 ; Search time 129.5 Seconds
(without alignments)
301.980 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 cttccctgcctccagccacacg 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 1026780

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/pdata/1/ina/1-COMB.seq:*
2: /cgn2_6/pdata/1/ina/5-COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A-COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B-COMB.seq:*
5: /cgn2_6/pdata/1/ina/H-COMB.seq:*
6: /cgn2_6/pdata/1/ina/PCITUS-COMB.seq:*
7: /cgn2_6/pdata/1/ina/PP-COMB.seq:*
8: /cgn2_6/pdata/1/ina/RE-COMB.seq:*
9: /cgn2_6/pdata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	13.6	61.8	27	2	US-08-460-890A-19 Sequence 19, Appl
C 2	13.6	61.8	27	2	US-08-460-890A-22 Sequence 22, Appl
C 3	13.6	61.8	27	3	US-08-167-641C-19 Sequence 19, Appl
C 4	13.6	61.8	27	3	US-08-167-641C-32 Sequence 22, Appl
C 5	13.6	61.8	27	3	US-08-460-971A-19 Sequence 19, Appl
C 6	13.6	61.8	27	3	US-08-460-971A-22 Sequence 22, Appl
C 7	13.6	61.8	27	3	US-08-462-040-19 Sequence 19, Appl
C 8	13.6	61.8	27	3	US-08-462-040-22 Sequence 22, Appl
C 9	13.4	60.9	25	3	US-09-396-196G-58648 Sequence 58648, A
C 10	13.4	60.9	25	3	US-09-396-196G-58649 Sequence 58649, A
C 11	13.4	60.9	25	3	US-09-396-196G-58650 Sequence 58650, A
C 12	13.4	59.1	23	2	US-08-766-982-10 Sequence 10, Appl
C 13	13.4	59.1	23	2	US-09-296-219-10 Sequence 10, Appl
C 14	13.4	59.1	24	2	US-08-529-190B-16 Sequence 16, Appl
C 15	13.4	59.1	30	3	US-09-119-507B-10 Sequence 10, Appl
C 16	13.4	59.1	30	3	US-09-119-507B-18 Sequence 18, Appl
C 17	13.4	59.1	30	3	US-08-897-556A-38 Sequence 38, Appl
C 18	13.4	59.1	30	3	US-08-897-556A-10 Sequence 10, Appl
C 19	13.4	59.1	30	3	US-09-547-693-10 Sequence 10, Appl
C 20	13.4	59.1	30	3	US-09-547-693-38 Sequence 38, Appl
C 21	12.8	58.2	17	3	US-08-944-410-88 Sequence 84, Appl
C 22	12.8	58.2	25	3	US-09-396-196G-98155 Sequence 98155, A
C 23	12.8	58.2	25	3	US-09-396-196G-98156 Sequence 98156, A
C 24	12.6	57.3	24	2	US-08-529-190B-14 Sequence 14, Appl

25	12.6	57.3	24	3	US-07-792-600-8 Sequence 8, Appl
26	12.6	57.3	24	3	US-09-157-021-8 Sequence 8, Appl
27	12.6	57.3	24	3	US-09-156-842-8 Sequence 8, Appl
28	12.6	57.3	24	3	US-09-591-514-8 Sequence 8, Appl
29	12.6	57.3	25	3	US-09-396-196G-2327 Sequence 2327, Ap
30	12.6	57.3	25	3	US-09-396-196G-44454 Sequence 44454, A
C 31	12.6	57.3	25	3	US-09-396-196G-76723 Sequence 76723, A
C 32	12.6	57.3	27	6	PCT-US93-02352-4 Sequence 4, Appl
C 33	12.6	57.3	29	6	PCT-US93-02352-6 Sequence 6, Appl
C 34	12.4	56.4	19	3	US-09-517-467B-9 Sequence 9, Appl
C 35	12.4	56.4	24	3	US-08-899-112B-17 Sequence 17, Appl
C 36	12.4	56.4	24	3	US-09-887-145-6 Sequence 6, Appl
C 37	12.4	56.4	24	3	US-09-011-553-18 Sequence 18, Appl
C 38	12.4	56.4	24	3	US-09-971-773-61 Sequence 61, Appl
C 39	12.4	56.4	25	2	US-08-268-020A-1 Sequence 1, Appl
C 40	12.4	56.4	25	3	US-08-899-112B-18 Sequence 18, Appl
C 41	12.4	56.4	25	3	US-09-011-553-19 Sequence 19, Appl
C 42	12.4	56.4	25	3	US-09-396-196G-43715 Sequence 43715, A
C 43	12.4	56.4	25	3	US-09-396-196G-43716 Sequence 43716, A
C 44	12.4	56.4	25	3	US-09-396-196G-56734 Sequence 56734, A
C 45	12.4	56.4	25	3	US-09-396-196G-61491 Sequence 61491, A

ALIGNMENTS

RESULT 1
US-08-460-890A-19/c
Sequence 19, Application US/08460890A
Patent No. 5994109
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchaik, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: PARCISO for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,890A
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Waipury, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 488-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:

LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-460-890A-19

Query Match 61.8%; Score 13.6; DB 2; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 20 CTCCTCGCCTCCCTCCCA 1

RESULT 2
US-08-460-890A-22/c
Sequence 22, Application US/08460890A

PATENT No. 5994109
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB

OPERATING SYSTEM: IBM Compatible
SOFTWARE: FASTSEQ for Windows 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,890A
FILING DATE: June 5, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389

FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/066
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-460-890A-22

Query Match 61.8%; Score 13.6; DB 2; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 20 CTCCTCGCCTCCCTCCCA 1

RESULT 3
US-08-167-641C-19/c
Sequence 19, Application US/08167641C
Patent No. 6033884

GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB

OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/167,641C
FILING DATE: December 14, 1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725

FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 205/012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440

TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:

LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-167-641C-19

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 20 CTCCTCGCCTCCCTCCCA 1

RESULT 4
US-08-167-641C-22/c
Sequence 22, Application US/08167641C
Patent No. 6033884
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.

```

; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEEX: 67-3510
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-167-641C-22

Query Match      61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 CTCCTCGCCTCCAGCCACA 20
Db      20 CTCCTCGCCTCCCTCCCA 1

RESULT 5
US-08-460-971A-19/C
; Sequence 19, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
```

```

; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,971A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEEX: 67-3510
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-460-971A-19

Query Match      61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 CTCCTCGCCTCCAGCCACA 20
Db      20 CTCCTCGCCTCCCTCCCA 1

RESULT 6
US-08-460-971A-22/C
; Sequence 22, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,971A
```

FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-460-971A-22

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCTCCAGCCACA 20
DB 20 CTCCTCGCTCCAGCCACA 1

RESULT 7
US-08-462-040-19/c
Sequence 19, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: Moo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gotchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,040
FILING DATE: June 5, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-462-040-19

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCTCCAGCCACA 20
DB 20 CTCCTCGCTCCAGCCACA 1

RESULT 8
US-08-462-040-22/c
Sequence 22, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: Moo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gotchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,040
FILING DATE: June 5, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-462-040-22

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCCCTCCGACCCACA 20
DB 20 CTCCTCCCTCCCTCCCA 1

RESULT 9
US-09-396-196G-58648/c
Sequence 58648, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
PRIORITY FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58648
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58648

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 24 GCCTCCAGCCACAG 10

RESULT 10
US-09-396-196G-58649/c
Sequence 58649, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Mack
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
PRIORITY FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58649
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58649

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 18 GCCTCCAGCCACAG 4

RESULT 11
US-09-396-196G-58650/c
Sequence 58650, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Mack
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
PRIORITY FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58650
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58650

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 15 GCCTCCAGCCACAG 1

RESULT 12
US-08-766-982-10
Sequence 10, Application US/08766982
Patent No. 5948892
GENERAL INFORMATION:
APPLICANT: Wahl, Robert C.
TITLE OF INVENTION: Analogs of Macrophage Stimulating
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,982
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-441
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-766-982-10

Query Match 59.1%; Score 13; DB 2; Length 23;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCCCTCGCCTCCAGCCACATG 22

DB 2 TGCCTCTCCTCCTCCGACATATG 22

RESULT 13

US-09-296-219-10

Sequence 10, Application US/09296219

Patent No. 6248560

GENERAL INFORMATION:

APPLICANT: Wahl, Robert C.

TITLE OF INVENTION: Analogs of Macrophage Stimulating

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 Beverlyland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/296,219

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-441

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-296-219-10

Query Match

Best Local Similarity 59.1%; Score 13; DB 3; Length 23;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCCCTCGCCTCCAGCCACATG 22

DB 2 TGCCTCTCCTCCTCCGACATATG 22

RESULT 14

US-08-529-190B-16

Sequence 16, Application US/08529190B

Patent No. 5833991

GENERAL INFORMATION:

APPLICANT: Masucci, Maria G.

TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: One Financial Center

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/529,190B

FILING DATE: 15-SEP-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE9501324-9

FILING DATE: 10-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US08/522,595

FILING DATE: 01-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Ph.D., Kathleen A

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3255/53015

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 24 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

US-08-529-190B-16

Query Match 59.1%; Score 13; DB 2; Length 24;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACAT 21

DB 4 CACCCGACCTCCAGCTCCAT 24

RESULT 15

US-09-119-507B-10

Sequence 10, Application US/09119507B

Patent No. 6548642

GENERAL INFORMATION:

APPLICANT: Kieliszewski, Marcia J.

TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums

FILE REFERENCE: OHU-03417

CURRENT APPLICATION NUMBER: US/09/119,507B

CURRENT FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 118

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO: 10

LENGTH: 30

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-119-507B-10

Query Match 59.1%; Score 13; DB 3; Length 30;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACAT 21

DB 5 CACCTTCACCTCCACCCCAT 25

Search completed: March 25, 2006, 13:03:00
Job time : 129.5 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).



This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 13:59:41 ; Search time 660.5 Seconds
(without alignments)
275.437 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22

Sequence: 1 CTCCTCGCCTCCAGCCACATG 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 10535742

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	US-10-798-652-7	Sequence 7, Appli
2	20.4	92.7	22	US-10-798-652-6	Sequence 6, Appli
3	15.8	71.8	25	US-11-036-317-589884	Sequence 589884,
4	15.6	70.9	25	US-11-036-317-179644	Sequence 179644,
5	15.6	70.9	25	US-11-036-317-250940	Sequence 250940,
6	15.6	70.9	25	US-11-036-317-253919	Sequence 253919,
7	15.6	70.9	25	US-11-036-317-268145	Sequence 268145,
8	15.6	70.9	25	US-11-036-317-362146	Sequence 362146,
9	15.6	70.9	25	US-11-036-317-391184	Sequence 391184,
10	15.6	70.9	25	US-11-060-756-199930	Sequence 199930,
11	15.4	70.0	20	US-10-994-138-20	Sequence 20, Appli
12	15.2	69.1	25	US-11-036-317-278200	Sequence 278200,
13	15.2	69.1	25	US-11-036-317-386177	Sequence 386177,
14	14.6	66.4	25	US-11-036-317-280075	Sequence 280075,
15	14.6	66.4	25	US-11-036-317-360952	Sequence 360952,
16	14.6	66.4	25	US-11-036-317-469308	Sequence 469308,
17	14.6	66.4	25	US-11-036-317-674241	Sequence 674241,
18	14.6	66.4	25	US-11-060-756-20600	Sequence 20600, A
19	14.6	66.4	25	US-11-060-756-20618	Sequence 20618, A
20	14.4	65.5	25	US-11-036-317-889096	Sequence 889096,
21	14.4	65.5	25	US-11-060-756-236805	Sequence 236805,
22	14.4	65.5	25	US-11-060-756-236806	Sequence 236806,
23	14.2	64.5	25	US-10-719-900-693282	Sequence 693282,

C 24	14.2	64.5	25	9	US-10-956-157-212800	Sequence 212800,
C 25	14.2	64.5	25	9	US-10-956-157-252303	Sequence 252303,
C 26	14.2	64.5	25	9	US-10-956-157-257894	Sequence 257894,
C 27	14.2	64.5	25	9	US-10-956-157-257895	Sequence 257895,
C 28	14.2	64.5	25	10	US-11-036-317-70925	Sequence 70925, A
C 29	14.2	64.5	25	10	US-11-036-317-343201	Sequence 343201, A
C 30	14.2	64.5	25	10	US-11-036-317-589885	Sequence 589885,
C 31	14.2	64.5	25	10	US-11-036-317-846119	Sequence 846119,
C 32	14	63.6	25	7	US-10-681-773-22238	Sequence 22238, A
C 33	14	63.6	25	7	US-10-681-773-72893	Sequence 72893, A
C 34	14	63.6	25	9	US-10-485-048-86	Sequence 86, Appli
C 35	14	63.6	25	9	US-10-956-157-174568	Sequence 174568,
C 36	14	63.6	25	9	US-10-956-157-250478	Sequence 250478,
C 37	14	63.6	25	10	US-11-036-317-516736	Sequence 516736,
C 38	14	63.6	25	10	US-11-036-317-923714	Sequence 923714,
C 39	13.8	62.7	24	3	US-09-940-185-956	Sequence 956, App
C 40	13.8	62.7	25	8	US-10-719-900-16414	Sequence 16414, A
C 41	13.8	62.7	25	9	US-10-956-157-163759	Sequence 163759,
C 42	13.8	62.7	25	10	US-11-036-317-516615	Sequence 516615,
C 43	13.8	62.7	25	10	US-11-036-317-579814	Sequence 579814,
C 44	13.8	62.7	25	10	US-11-036-317-579816	Sequence 579816,
C 45	13.6	61.8	25	5	US-10-215-112-12019	Sequence 12019, A

ALIGNMENTS

RESULT 1
US-10-798-652-7
Sequence 7, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-PO3136US01
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe
US-10-798-652-7

Query Match 100.0%; Score 22; DB 8; Length 22;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 1 CTCCTCGCCTCCAGCCACATG 22

RESULT 2
US-10-798-652-6
Sequence 6, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-PO3136US01
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7

```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 22
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: probe
US-10-798-652-6

Query Match          92.7%; Score 20.4; DB 8; Length 22;
Best Local Similarity 95.5%; Pred. No. 57;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 1 CTCCTCCACCTCCAGCCACATG 22

RESULT 3
US-11-036-317-589884/c
/ Sequence 589884, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 589884
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-589884

Query Match          71.8%; Score 15.8; DB 10; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.8e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCCTCGCCTCCAGCCACAT 21
Db 25 CCCTCGCCTCCAGCCCTCT 7

RESULT 4
US-11-036-317-179644/c
/ Sequence 179644, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 179644
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-179644

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 23 CCCACTCCCTCCAGCCACAG 2

RESULT 5
US-11-036-317-250940
/ Sequence 250940, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 250940
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-250940

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 2 CTCCTCCACATGAGCCACATG 23

RESULT 6
US-11-036-317-253919/c
/ Sequence 253919, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 253919
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-253919

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 22 CCCACTCCCTCCAGCCACAG 1

RESULT 7
US-11-036-317-268615
/ Sequence 268615, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
```

FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 268615
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-268615

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGGCTCCAGCCACATG 22
Db 1 CTCCTCCCATGCAGCCACCTG 22

RESULT 8
US-11-036-317-362146
Sequence 362146, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan
APPLICANT: Blume, John
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 362146
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-362146

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGGCTCCAGCCACATG 22
Db 3 CTCCTCCCATGCAGCCACCTG 24

RESULT 9
US-11-036-317-391184
Sequence 391184, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan
APPLICANT: Blume, John
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 391184
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-391184

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGGCTCCAGCCACATG 22
Db 4 CTCCTCCCATGCAGCCACCTG 25

RESULT 10
US-11-060-756-199930/c
Sequence 199930, Application US/11060756
Publication No. US20050221354A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mounts, William Martin
TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
FILE REFERENCE: AM101083 (031896-042000)
CURRENT APPLICATION NUMBER: US/11/060,756
CURRENT FILING DATE: 2005-02-18
NUMBER OF SEQ ID NOS: 303284
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 199930
LENGTH: 25
TYPE: DNA
ORGANISM: probe
US-11-060-756-199930

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGGCTCCAGCCACATG 22
Db 25 CTCCTCGGCTCCAGCCACATG 4

RESULT 11
US-10-994-138-20
Sequence 20, Application US/10994138
Publication No. US2005018666A1
GENERAL INFORMATION:
APPLICANT: Dow Global Technologies, Inc.
TITLE OF INVENTION: Improved Protein Expression Systems
FILE REFERENCE: 00588.105020
CURRENT APPLICATION NUMBER: US/10/994,138
CURRENT FILING DATE: 2004-11-19
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 20
LENGTH: 20
TYPE: DNA
ORGANISM: artificial
FEATURE:
OTHER INFORMATION: pyranalr
US-10-994-138-20

Query Match
Best Local Similarity 94.1%; Score 15.4; DB 9; Length 20;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CTCGCTTCAGCCACAT 21
Db 1 CTCGCTTCAGCCACAT 17

RESULT 12
US-11-036-317-278200/c
Sequence 278200, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan

```
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 278200
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-278200

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.6e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGGCTCCAGCCACA 20
Db      21 CCCACTCCCTCCAGCCACA 2

RESULT 13
US-11-036-317-386177/c
/ Sequence 386177, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 386177
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-386177

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.6e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGGCTCCAGCCACA 20
Db      20 CCCACTCCCTCCAGCCACA 1

RESULT 14
US-11-036-317-280075
/ Sequence 280075, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 280075
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
```

```
US-11-036-317-280075

Query Match          66.4%; Score 14.6; DB 10; Length 25;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGGCTCCAGCCACATG 21
Db      5 CTCCTCCCATGACAGCCACCT 25

RESULT 15
US-11-036-317-360952
/ Sequence 360952, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 360952
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-360952

Query Match          66.4%; Score 14.6; DB 10; Length 25;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      2 TCCCTCGGCTCCAGCCACATG 22
Db      1 TCCCTCCCATGACAGCCACCTG 21

Search completed: March 25, 2006, 15:50:32
Job time : 661.5 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 14:11:31 ; Search time 315.5 Seconds
(without alignments)
277.966 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 cccctcgcgcgcacacacatg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9258654 seqs, 1993127192 residues

Total number of hits satisfying chosen parameters: 13041728

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /SIDS5/pcodaca/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/pcodaca/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/pcodaca/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/pcodaca/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/pcodaca/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/pcodaca/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq1:*
9: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq2:*
13: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq3:*
14: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/pcodaca/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.8	71.8	20	8	US-10-310-914A-613434 Sequence 613434
2	15.8	71.8	21	8	US-10-310-914A-148615 Sequence 148615
3	15.8	71.8	23	8	US-10-310-914A-743419 Sequence 743419
4	15.4	70.0	22	8	US-10-310-914A-743458 Sequence 743458
5	15.2	69.1	20	8	US-10-310-914A-290068 Sequence 290068
6	15.2	69.1	21	8	US-10-310-914A-177473 Sequence 177473
7	15.2	69.1	21	8	US-10-310-914A-613446 Sequence 613446
8	15.2	69.1	24	8	US-10-310-914A-233742 Sequence 233742
9	15.2	69.1	24	8	US-10-310-914A-549830 Sequence 549830
10	15.2	69.1	25	8	US-10-310-914A-549844 Sequence 549844
11	15.2	69.1	22	8	US-10-310-914A-165922 Sequence 165922
12	14.8	67.3	18	8	US-10-310-914A-213337 Sequence 213337
13	14.8	67.3	19	8	US-10-310-914A-120157 Sequence 120157
14	14.8	67.3	19	8	US-10-310-914A-213338 Sequence 213338
15	14.8	67.3	19	8	US-10-310-914A-668834 Sequence 668834
16	14.8	67.3	19	8	US-10-310-914A-668892 Sequence 668892
17	14.8	67.3	20	8	US-10-310-914A-213339 Sequence 213339
18	14.8	67.3	20	8	US-10-310-914A-317759 Sequence 317759

C 19	14.8	67.3	20	8	US-10-310-914A-377554	Sequence 377554
C 20	14.8	67.3	20	8	US-10-310-914A-701328	Sequence 701328
C 21	14.8	67.3	20	8	US-10-310-914A-1138593	Sequence 1138593
C 22	14.8	67.3	20	14	US-11-093-746A-14	Sequence 14, Appl
C 23	14.8	67.3	21	8	US-10-310-914A-1108644	Sequence 1108644
C 24	14.8	67.3	21	8	US-10-310-914A-1138594	Sequence 1138594
C 25	14.8	67.3	25	14	US-11-121-849-439265	Sequence 439265
C 26	14.6	66.4	24	8	US-10-310-914A-237510	Sequence 237510
C 27	14.6	66.4	25	8	US-10-310-914A-91862	Sequence 91862, A
C 28	14.6	66.4	25	8	US-10-310-914A-91863	Sequence 91863, A
C 29	14.6	66.4	29	8	US-10-857-780-2372	Sequence 2372, Ap
C 30	14.4	65.5	21	8	US-10-310-914A-532861	Sequence 532861
C 31	14.4	65.5	21	8	US-10-310-914A-102269	Sequence 102269
C 32	14.4	65.5	21	8	US-10-310-914A-102291	Sequence 102291
C 33	14.4	65.5	21	8	US-10-310-914A-532862	Sequence 532862
C 34	14.4	65.5	21	8	US-10-310-914A-782776	Sequence 782776
C 35	14.4	65.5	22	8	US-10-310-914A-782816	Sequence 782816
C 36	14.4	65.5	25	9	US-10-933-982-36071	Sequence 36071, A
C 37	14.2	64.5	19	8	US-10-310-914A-372336	Sequence 372336
C 38	14.2	64.5	19	8	US-10-310-914A-654165	Sequence 654165
C 39	14.2	64.5	20	8	US-10-310-914A-448406	Sequence 448406
C 40	14.2	64.5	20	8	US-10-310-914A-519040	Sequence 519040
C 41	14.2	64.5	20	8	US-10-310-914A-1002041	Sequence 1002041
C 42	14.2	64.5	20	8	US-10-310-914A-1223911	Sequence 1223911
C 43	14.2	64.5	20	8	US-10-310-914A-1352507	Sequence 1352507
C 44	14.2	64.5	21	8	US-10-310-914A-724597	Sequence 724597
C 45	14.2	64.5	21	8	US-10-310-914A-1352505	Sequence 1352505

ALIGNMENTS

RESULT 1
US-10-310-914A-613434
; Sequence 613434, Application US/10310914A
; Publication No. US2006003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Biologically detectable group of novel regulatory genes an
; FILE REFERENCE: 06087, 0200, CPUS01
; CURRENT APPLICATION NUMBER: US/10/310, 914A
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 613434
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-613434

Query Match 71.8%, Score 15.8; DB 8; Length 20;
Best Local Similarity 73.7%; Pred. No. 1.9e+03;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCCTCGCTCCAGCCACAT 21
DB 2 CCCTGACCCGACGACCAU 20

RESULT 2
US-10-310-914A-148615
; Sequence 148615, Application US/10310914A
; Publication No. US2006003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Biologically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087, 0200, CPUS01
; CURRENT APPLICATION NUMBER: US/10/310, 914A
; CURRENT FILING DATE: 2002-12-06

NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 148615
LENGTH: 21
TYPE: RNA
ORGANISM: Human
US-10-310-914A-148615

Query Match 71.8%; Score 15.8; DB 8; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCAC 19
|:|||||:|||||
Db 1 CUCGCCGCCGCCAGCCUC 19

RESULT 3
US-10-310-914A-743419/c
Sequence 743419, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:

APPLICANT: Bentwich, Isaac
APPLICANT: Shiler, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 743419
LENGTH: 23
TYPE: RNA
ORGANISM: Human
US-10-310-914A-743419

Query Match 71.8%; Score 15.8; DB 8; Length 23;
Best Local Similarity 89.5%; Pred. No. 1.9e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCAC 19
|:|||||:|||||
Db 22 CTCCTCGGCCGCCAGCCCC 4

RESULT 4
US-10-310-914A-743458/c
Sequence 743458, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:

APPLICANT: Bentwich, Isaac
APPLICANT: Shiler, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 743458
LENGTH: 22
TYPE: RNA
ORGANISM: Human
US-10-310-914A-743458

Query Match 70.0%; Score 15.4; DB 8; Length 22;
Best Local Similarity 94.1%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCC 17
|:|||||:|||||
Db 18 CTCCTCGGCCGCCAGCC 2

RESULT 5
US-10-310-914A-290068
Sequence 290068, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:

APPLICANT: Bentwich, Isaac
APPLICANT: Shiler, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 290068
LENGTH: 20
TYPE: RNA
ORGANISM: Human
US-10-310-914A-290068

Query Match 69.1%; Score 15.2; DB 8; Length 20;
Best Local Similarity 70.0%; Pred. No. 3.4e+03;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACA 20
|:|||||:|||||
Db 1 CUCGCCGCCGCCAGCCUCA 20

RESULT 6
US-10-310-914A-177473
Sequence 177473, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:

APPLICANT: Bentwich, Isaac
APPLICANT: Shiler, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 177473
LENGTH: 21
TYPE: RNA
ORGANISM: Human
US-10-310-914A-177473

Query Match 69.1%; Score 15.2; DB 8; Length 21;
Best Local Similarity 80.0%; Pred. No. 3.4e+03;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACA 20
|:|||||:|||||
Db 1 CUCGCCGCCGCCAGCCUCA 20

RESULT 7
US-10-310-914A-631446
Sequence 631446, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:

APPLICANT: Bentwich, Isaac
APPLICANT: Shiler, Kyuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402

```
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 631446
/ LENGTH: 21
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-631446
```

```
Query Match
Best Local Similarity 69.1%; Score 15.2; DB 8; Length 21;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 TCCTCGCCTCCAGCCACAT 21
Db 2 UCCUCGCGCCCGCCAGCCACAU 21
```

```
RESULT 8
US-10-310-914A-233742
/ Sequence 233742, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ CURRENT FILING DATE: 2002-12-06
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 233742
/ LENGTH: 24
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-233742
```

```
Query Match
Best Local Similarity 75.0%; Score 15.2; DB 8; Length 24;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 CCCTCGCCTCCAGCCACATG 22
Db 1 CCCCAACCCGCCAGCCACAU 20
```

```
RESULT 9
US-10-310-914A-549830
/ Sequence 549830, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ CURRENT FILING DATE: 2002-12-06
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 549830
/ LENGTH: 24
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-549830
```

```
Query Match
Best Local Similarity 69.1%; Score 15.2; DB 8; Length 24;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCGCCTCCAGCCACA 20
Db 1 CUCCUCGCGCCCGCCAGCCCA 20
```

```
RESULT 10
US-10-310-914A-549844
/ Sequence 549844, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ CURRENT FILING DATE: 2002-12-06
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 549844
/ LENGTH: 25
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-549844
```

```
Query Match
Best Local Similarity 69.1%; Score 15.2; DB 8; Length 25;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCGCCTCCAGCCACA 20
Db 4 CUCCUCGCGCCCGCCAGCCCA 23
```

```
RESULT 11
US-10-310-914A-165922/C
/ Sequence 165922, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ CURRENT FILING DATE: 2002-12-06
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 165922
/ LENGTH: 22
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-165922
```

```
Query Match
Best Local Similarity 68.2%; Score 15; DB 8; Length 22;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCGCCTCCAG 15
Db 22 CTCCTCGCCTCCAG 8
```

```
RESULT 12
US-10-310-914A-21337/C
/ Sequence 21337, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ CURRENT FILING DATE: 2002-12-06
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
```

SEQ ID NO 213337
LENGTH: 18
TYPE: RNA
ORGANISM: Human
US-10-310-914A-213337

Query Match 67.3%; Score 14.8; DB 8; Length 18;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCA 18
DB 18 CTCCTCGCCTCCAGTCA 1

RESULT 13
US-10-310-914A-120157/c
Sequence 120157, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 120157
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-120157

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCTCGCCTCCAGCCAC 19
DB 19 TCCTCGCCTCCAGCCAC 2

RESULT 14
US-10-310-914A-213338/c
Sequence 213338, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 213338
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-213338

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCA 18
DB 18 CTCCTCGCCTCCAGTCA 1

RESULT 15
US-10-310-914A-668834/c
Sequence 668834, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 668834
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-668834

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCTCGCCTCCAGCCACA 20
DB 18 CCTCGCCTCCAGCCACA 1

Search completed: March 25, 2006, 14:54:05
Job time : 316.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:28:17 ; Search time 440.5 Seconds
(without alignments)
5103.854 Million cell updates/sec

Title: SEQ1-69C

Perfect score: 564
Sequence: 1 gcagccctgcctcagaac.....agtgcgcttcgcgactc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9258654 seqs, 1993127192 residues

Total number of hits satisfying chosen parameters: 18517308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database : Published Applications NA New:
1: /SIDS/ptodata/1/pubpna/US08_NEW_PUB.seq:
2: /SIDS/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /SIDS/ptodata/1/pubpna/US07_NEW_PUB.seq:
4: /SIDS/ptodata/1/pubpna/PCT_NEW_PUB.seq:
5: /SIDS/ptodata/1/pubpna/US09_NEW_PUB.seq:
6: /SIDS/ptodata/1/pubpna/US09_NEW_PUB.seq:
7: /SIDS/ptodata/1/pubpna/US10_NEW_PUB.seq:
8: /SIDS/ptodata/1/pubpna/US10_NEW_PUB.seq:
9: /SIDS/ptodata/1/pubpna/US10_NEW_PUB.seq:
10: /SIDS/ptodata/1/pubpna/US10_NEW_PUB.seq:
11: /SIDS/ptodata/1/pubpna/US11_NEW_PUB.seq:
12: /SIDS/ptodata/1/pubpna/US11_NEW_PUB.seq:
13: /SIDS/ptodata/1/pubpna/US11_NEW_PUB.seq:
14: /SIDS/ptodata/1/pubpna/US11_NEW_PUB.seq:
15: /SIDS/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	553	98.0	169495	14 US-11-121-086-61	Sequence 61, Appl
2	39.8	7.1	541	6 US-09-925-065A-734471	Sequence 734471, A
3	38	6.7	162289	14 US-11-121-086-20	Sequence 20, Appl
4	37	6.6	712	8 US-10-750-185-32987	Sequence 32987, A
5	36.4	6.5	1172	8 US-09-925-065A-57042	Sequence 57042, A
6	36.4	6.5	1172	9 US-10-301-480-158280	Sequence 158280, A
7	36.4	6.5	1172	10 US-10-301-480-771689	Sequence 771689, A
8	35.8	6.3	201	8 US-10-995-561-63329	Sequence 63329, A
9	35.8	6.3	201	8 US-10-995-561-69610	Sequence 69610, A
10	35.8	6.3	25257	8 US-10-995-561-13412	Sequence 13412, A
11	35.8	6.3	86361	8 US-10-995-561-13364	Sequence 13364, A
12	35.6	6.3	179597	14 US-11-121-086-91	Sequence 91, Appl
13	35	6.2	1741	9 US-10-301-480-93596	Sequence 93596, A
14	35	6.2	1741	10 US-10-301-480-707005	Sequence 707005, A
15	34.8	6.2	835	8 US-10-750-185-42100	Sequence 42100, A
16	34.8	6.2	835	8 US-10-750-185-42100	Sequence 42100, A
17	34.8	6.2	1198	6 US-09-925-065A-42328	Sequence 42328, A
18	34.8	6.2	1198	6 US-09-925-065A-42328	Sequence 42328, A

19	34.8	6.2	1198	6 US-09-925-065A-42329	Sequence 42329, A
20	34.8	6.2	1198	9 US-10-301-480-143566	Sequence 143566, A
21	34.8	6.2	1198	9 US-10-301-480-143567	Sequence 143567, A
22	34.8	6.2	1198	10 US-10-301-480-756975	Sequence 756975, A
23	34.8	6.2	1198	10 US-10-301-480-756976	Sequence 756976, A
24	34.8	6.2	1533	6 US-09-925-065A-685510	Sequence 685510, A
25	34.8	6.2	1533	6 US-09-925-065A-685511	Sequence 685511, A
26	34.6	6.1	220895	8 US-10-775-169-88	Sequence 88, Appl
27	34.4	6.1	201	8 US-10-995-561-12574	Sequence 12574, A
28	34.4	6.1	201	8 US-10-995-561-12594	Sequence 12594, A
29	34.4	6.1	1926	9 US-10-301-480-88052	Sequence 88052, A
30	34.4	6.1	1926	10 US-10-301-480-701461	Sequence 701461, A
31	34.4	6.1	5051	8 US-10-995-561-493	Sequence 493, App
32	34.4	6.1	5232	8 US-10-995-561-492	Sequence 492, App
33	34	6.0	699	6 US-09-925-065A-918774	Sequence 918774, A
34	34	6.0	811	6 US-09-925-065A-919302	Sequence 919302, A
35	34	6.0	813	6 US-09-925-065A-917128	Sequence 917128, A
36	34	6.0	169725	14 US-11-121-086-63	Sequence 63, Appl
37	34	6.0	200628	14 US-11-121-086-62	Sequence 62, Appl
38	33.8	6.0	527	6 US-09-925-065A-736105	Sequence 736105, A
39	33.8	6.0	527	6 US-09-925-065A-736106	Sequence 736106, A
40	33.8	6.0	527	6 US-09-925-065A-813895	Sequence 813895, A
41	33.8	6.0	665	6 US-09-925-065A-726390	Sequence 726390, A
42	33.4	5.9	2508	11 US-11-072-512-1958	Sequence 1958, Ap
43	33.2	5.9	509	6 US-09-925-065A-248642	Sequence 248642, A
44	33.2	5.9	509	6 US-09-925-065A-248643	Sequence 248643, A
45	33.2	5.9	514	10 US-10-301-480-328622	Sequence 328622, A

ALIGNMENTS

RESULT 1
US-11-121-086-61/c
Sequence 61, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
PRIOR FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 61
LENGTH: 169495
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (70072)..(70171)
OTHER INFORMATION: a, c, g, t, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (139457)..(157244)
OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61

Query Match 98.0%; Score 553; DB 138; Length 169495;

Best Local Similarity 99.8%; Pred. No. 1,4e-138; Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GCAGCCTGCTCAGAAACGAGAGCAGCAGCACTCAGGTGACTCAGCCCATGTG 60
Db 51758 GCAGCCTGCTCAGAAACGAGAGCAGCAGCACTCAGGTGACTCAGCCCATGTG 51699
Qy 61 GCTGAGAGCGAGAGCCTCTCTAGAGCAGGCGCAGGCGCAGCTCAGGTGAGCGCA 120
Db 51698 GCTGAGAGCGAGAGCCTCTCTAGAGCAGGCGCAGGCGCAGCTCAGGTGAGCGCA 51639

QY	121	GGGGCTTTGGCAGTGGGCAAGGGGGCTTCATACAGCTTACTCACTGACATCGAGTCC	180
Db	51638	GGGGCTTTGGCAGTGGGCAAGGGGGCTTCATACAGCTTACTCACTGACATCGAGTCC	51579
QY	181	CTGGTGCACGCTCTTGAAATGTGAAATGTGAAATGTTTCCATTAAAGAAAGTGTGTG	240
Db	51578	CTGGTGCACGCTCTTGAAATGTGAAATGTGAAATGTTTCCATTAAAGAAAGTGTGTG	51519
QY	241	GCGGGCCATGCCCCCAGAGTTGACACACTGACCTGTCAGAGGGTTGGGGCTTCCAGTC	300
Db	51518	GCGGGCCATGCCCCCAGAGTTGACACACTGACCTGTCAGAGGGTTGGGGCTTCCAGTC	51459
QY	301	ACAGGGTCCCATCCACGTACAGAGCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAGTAA	360
Db	51458	ACAGGGTCCCATCCACGTACAGAGCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAGTAA	51399
QY	361	CCAAAGGAGGCTTGGGAAACCAATCTGAAAGGGCATGGCTTTGATTATTAAGAGGGGTG	420
Db	51398	CCAAAGGAGGCTTGGGAAACCAATCTGAAAGGGCATGGCTTTGATTATTAAGAGGGGTG	51339
QY	421	GGCTGGGCTGGGCAAGGCCACAGGTCTGAGTCAGAGCCAGAGGCAAGAGCTGTGCCCC	480
Db	51338	GGCTGGGCTGGGCAAGGCCACAGGTCTGAGTCAGAGCCAGAGGCAAGAGCTGTGCCCC	51279
QY	481	A-GCACTGCCGCGCGCTCTGCAGTCAGTCTCTTGGCCACTGAGAAACAGCTGTAGA	539
Db	51278	AGGCACTGCCGCGCGCTCTGCAGTCAGTCTCTTGGCCACTGAGAAACAGCTGTAGA	51219
QY	540	GAGGCAATGAGGCTTTTCCGACATTC	564
Db	51218	GAGGCAATGAGGCTTTTCCGACATTC	51194

```

RESULT 2
US-09-925-065A-734471
: Sequence 734471, Application US/09925065A
: Publication No. US20040181048A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single
: TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.135
: CURRENT APPLICATION NUMBER: US/09/925, 065A
: PRIOR FILING DATE: 2001-08-08
: PRIOR APPLICATION NUMBER: US 60/243, 096
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: US 60/252, 147
: PRIOR FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: US 60/250, 092
: PRIOR FILING DATE: 2000-11-30
: PRIOR APPLICATION NUMBER: US 60/261, 766
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/289, 846
: PRIOR FILING DATE: 2001-05-09
: NUMBER OF SEQ ID NOS: 957086
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 734471
: LENGTH: 541
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-065A-734471

```

Query Match	7.1%	Score 39.8	DB 6	Length 541
Best Local Similarity	64.8%	Pred. No. 0.64		
Matches	59	Conservative	0	Mismatches 32
			Indels 0	Gaps 0

QY	4	GCCCTGCTCGAAGAAACAGAGAGACGCACACATCAGGTGATCATCACCCTCATGTGCT	63
DB	442	GCCCTCCTCGAAGAACAGAAAGCTGGGACCTCTCGTGGGCTCACCTCATTTGGAA	501

QY	64	GGAGCGGAGAGAGCTCTCTGAGGACAGGSCCA	94
DB	502	TGATGAGAGGTGTTCCCAAGGCTCTGGGSCCA	532

```

RESULT 3
US-11-121-086-20/C
; Sequence 20, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-0000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 162289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-20

```

Query Match	6.7%	Score 18	DB 14	Length 16289	
Best Local Similarity	46.6%	Pred. No. 5.2	Mismatches 140	Indels 0	Gaps 0
Matches 122	Conservative 0				
QY	3	AGCCCTGCTCAGAAAAACAGAGACGACGACACTCAGCGTGACTCACCCCCCATGTGGC	62		
Db	145401	AGCTCACACAGCAACCTGGGGAGGCGACCTCTGTGGCTTGGAGTGGGATACAGAAAG	145342		
QY	63	TGAGAGGAGGAGGACCTCTGAGGACAGGGCCAGGCGACCGTCAGTGGGTGACGCGCAG	122		
Db	145341	CCTTAGTCACAGGAGGAGACTGGAAGCCAGGGCCTCAAGAGAGTGGGGTGGGCATTGGGGAG	145282		
QY	123	GGTCTTGCCATGTGTGGGCACAGGGCTTGCAATACGCTTACTGATGACATGAGATCCCT	182		
Db	145381	GTCATGGGCAAGTGGGCGAGGTCAAGCTGTGAGCACTTCCCGAGGCCAATACCAAAAT	145222		
QY	183	GGTCCAGACCTCTGGAAGTGTGAAAGTAGCAATTTTCCATTAAAGAAAGTGTGTGC	242		
Db	145221	GAGACAACTTTGGGACTGACAGAGCATGTGGGGGCTCTGGAGGACGAGAAAGTAAATGTTC	145162		
QY	243	CGGCCATGCCCCCCCAAGCTTGC	264		
Db	145161	ACACAAACACCCCAAGCTTGC	145140		

```

RESULT 4
US-10-750-185-32987
? Sequence 32987, Application US/10750185
? Publication No. US20050260603A1
? GENERAL INFORMATION:
? APPLICANT: MMI GENOMICS, INC.
? APPLICANT: DENISE, Sue K.
? APPLICANT: KERR, Richard
? APPLICANT: ROSENFELD, David
? APPLICANT: HOLM, Tom
? APPLICANT: BATES, Stephen
? APPLICANT: FAVTIN, Dennis
? TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
? FILE REFERENCE: MM11100-2
? CURRENT APPLICATION NUMBER: US/10/750.185
? CURRENT FILING DATE: 2003-12-31
? PRIOR APPLICATION NUMBER: US 60/437,482
? PRIOR FILING DATE: 2002-12-31
? NUMBER OF SEQ ID NOS: 64922
? SOFTWARE: PatentIN version 3.1
? SEQ ID NO 32987
? LENGTH: 712
? TYPE: DNA
? ORGANISM: Bovine 19866880739882
? US-10-750-185-32987

```

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 3.8;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
OY 307 TCCATCCACGATACCAAGCCAGGCTGCGAGAGGTCCTCCGCACTGATGAACCAAG 366
DB 283 TGCCGACACATTCAGTCTCTATGGAAAGATGAATAATTAAGAGATGAATAATGCTG 342
OY 367 GAGGCTGGGAAACCACTGTAAGGAGATGCTTGAATTAAGAGAGGCTGGGCTGG 426
DB 343 AAAAAGTGTAAAGACATTAATTAAGCATTAATGCTTACTGTGAAGAGAGGCTCT 402
OY 427 GCTGGGCAAGGCCACAGGCTGAGTCAAGAGCCAGAGGAGAGGCTGCTCCAGCACT 486
DB 403 TGCTTACTCCCTTACTGGAATAATAGCGAAGGCCAGGATGATCCAGGATCCTCAAGCT 462
OY 487 GCGGCGGCTCTGCGATGAGTCTCTCGGCCACCTGAGAACAGCCTG 535
DB 463 TGCTCATCTCTGTCTTCCATTTCTCAATTAAGAGCCTCGAAGAGACTG 511
```

RESULT 5
US-10-750-623-32987
Sequence 32987, Application US/10750623

GENERAL INFORMATION:
APPLICANT: MMR GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OR INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT FILING DATE: 2003-12-31
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32987
LENGTH: 712
TYPE: DNA
ORGANISM: Bovine 19866880739882
US-10-750-623-32987

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 3.8;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
OY 307 TCCATCCACGATACCAAGCCAGGCTGCGAGAGGTCCTCCGCACTGATGAACCAAG 366
DB 283 TGCCGACACATTCAGTCTCTATGGAAAGATGAATAATTAAGAGATGAATAATGCTG 342
OY 367 GAGGCTGGGAAACCACTGTAAGGAGATGCTTGAATTAAGAGAGGCTGGGCTGG 426
DB 343 AAAAAGTGTAAAGACATTAATTAAGCATTAATGCTTACTGTGAAGAGAGGCTCT 402
OY 427 GCTGGGCAAGGCCACAGGCTGAGTCAAGAGCCAGAGGAGAGGCTGCTCCAGCACT 486
DB 403 TGCTTACTCCCTTACTGGAATAATAGCGAAGGCCAGGATGATCCAGGATCCTCAAGCT 462
OY 487 GCGGCGGCTCTGCGATGAGTCTCTCGGCCACCTGAGAACAGCCTG 535
DB 463 TGCTCATCTCTGTCTTCCATTTCTCAATTAAGAGCCTCGAAGAGACTG 511
```

RESULT 6
US-09-925-065A-57042/c
Sequence 57042, Application US/09925065A
Publication No. US20040181048A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single
TITLE OR INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57042
LENGTH: 1172
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-57042

Query Match 6.5%; Score 36.4; DB 6; Length 1172;
Best Local Similarity 59.8%; Pred. No. 6;
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
OY 25 GAGCGACGACACTCAGGTCAGTCAACCCCATGTGCTGGAGGCGAGGAGCCTCTGA 84
DB 599 GGGGCGAGCCCTGCGATGTGATCTTTGGCAGATGAGAGAGACACAGAGCAGCAGA 540
OY 85 GGCAGGGCGAGGCGAGCCTGAGTGGGTGACGCGAGGGTC 126
DB 539 GGAAGGAGAGCGAGCCAGCAGCAGGATGACAGGTGACATC 498
```

RESULT 7
US-10-301-480-158280/c
Sequence 158280, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 158280
LENGTH: 1172
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-158280

Query Match 6.5%; Score 36.4; DB 9; Length 1172;
Best Local Similarity 59.8%; Pred. No. 6;
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
OY 25 GAGCGACGACACTCAGGTCAGTCAACCCCATGTGCTGGAGGCGAGGAGCCTCTGA 84
DB 599 GGGGCGAGCCCTGCGATGTGATCTTTGGCAGATGAGAGAGACACAGAGCAGCAGA 540
OY 85 GGCAGGGCGAGGCGAGCCTGAGTGGGTGACGCGAGGGTC 126
DB 539 GGAAGGAGAGCGAGCCAGCAGCAGGATGACAGGTGACATC 498
```


Db 70 GTATGACAGAGTGCATTTCGGTGAAGTGCCTCAGT 36

Search completed: March 25, 2006, 18:31:06
Job time : 441.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:28:17 ; Search time 440.5 Seconds
(without alignments)
5103.654 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaac.....agtcgctcttcgagcttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9258654 seqs, 1993127192 residues

Total number of hits satisfying chosen parameters: 18517308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA_New:*
1: /SID55/pcodata/1/pubpna/US06_NEW_PUB.seq:*
2: /SID55/pcodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SID55/pcodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SID55/pcodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SID55/pcodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SID55/pcodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SID55/pcodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /SID55/pcodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /SID55/pcodata/1/pubpna/US10_NEW_PUB.seq2:*
10: /SID55/pcodata/1/pubpna/US10_NEW_PUB.seq3:*
11: /SID55/pcodata/1/pubpna/US11_NEW_PUB.seq:*
12: /SID55/pcodata/1/pubpna/US11_NEW_PUB.seq:*
13: /SID55/pcodata/1/pubpna/US11_NEW_PUB.seq3:*
14: /SID55/pcodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SID55/pcodata/1/pubpna/US66_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	551.4	97.8	169495	14 US-11-121-086-61	Sequence 61, Appl
2	39.8	7.1	541	6 US-09-925-065A-734471	Sequence 734471, A
3	39.6	7.0	162289	14 US-11-121-086-20	Sequence 20, Appl
4	37	6.6	712	8 US-10-750-185-32987	Sequence 32987, A
5	37	6.6	712	8 US-10-750-623-32987	Sequence 32987, A
6	35.8	6.3	201	8 US-10-995-561-63329	Sequence 63329, A
7	35.8	6.3	201	8 US-10-995-561-69610	Sequence 69610, A
8	35.8	6.3	25257	8 US-10-995-561-13412	Sequence 13412, A
9	35.8	6.3	86361	8 US-10-995-561-13364	Sequence 13364, A
10	35.6	6.3	179597	14 US-11-121-086-91	Sequence 91, Appl
11	35	6.2	1741	9 US-10-301-480-93596	Sequence 93596, A
12	35	6.2	1741	10 US-10-301-480-70705	Sequence 70705, A
13	34.8	6.2	835	8 US-10-750-185-42100	Sequence 42100, A
14	34.8	6.2	835	8 US-10-750-623-42100	Sequence 42100, A
15	34.8	6.2	1172	6 US-09-925-065A-57042	Sequence 57042, A
16	34.8	6.2	1172	6 US-10-301-480-158280	Sequence 158280, A
17	34.8	6.2	1172	10 US-10-301-480-771689	Sequence 771689, A
18	34.8	6.2	1198	6 US-09-925-065A-42328	Sequence 42328, A

19	34.8	6.2	1198	6 US-09-925-065A-42329	Sequence 42329, A
20	34.8	6.2	1198	9 US-10-301-480-143566	Sequence 143566, A
21	34.8	6.2	1198	9 US-10-301-480-143567	Sequence 143567, A
22	34.8	6.2	1198	10 US-10-301-480-756975	Sequence 756975, A
23	34.8	6.2	1198	10 US-10-301-480-756976	Sequence 756976, A
24	34.8	6.2	1533	6 US-09-925-065A-685510	Sequence 685510, A
25	34.8	6.2	1533	6 US-09-925-065A-685511	Sequence 685511, A
26	34.6	6.1	220895	8 US-10-775-169-88	Sequence 88, Appl
27	34.4	6.1	201	8 US-10-995-561-12574	Sequence 12574, A
28	34.4	6.1	201	8 US-10-995-561-12594	Sequence 12594, A
29	34.4	6.1	1926	9 US-10-301-480-88052	Sequence 88052, A
30	34.4	6.1	1926	10 US-10-301-480-701461	Sequence 701461, A
31	34.4	6.1	5051	8 US-10-995-561-493	Sequence 493, Appl
32	34.4	6.1	5232	8 US-10-995-561-492	Sequence 492, Appl
33	34	6.0	699	6 US-09-925-065A-918774	Sequence 918774, A
34	34	6.0	811	6 US-09-925-065A-919302	Sequence 919302, A
35	34	6.0	813	6 US-09-925-065A-917128	Sequence 917128, A
36	34	6.0	169725	14 US-11-121-086-63	Sequence 63, Appl
37	34	6.0	200628	14 US-11-121-086-62	Sequence 62, Appl
38	33.8	6.0	527	6 US-09-925-065A-736105	Sequence 736105, A
39	33.8	6.0	527	6 US-09-925-065A-736106	Sequence 736106, A
40	33.8	6.0	527	6 US-09-925-065A-813895	Sequence 813895, A
41	33.8	6.0	665	6 US-09-925-065A-726390	Sequence 726390, A
42	33.4	5.9	2508	11 US-11-072-512-1958	Sequence 1958, Ap
43	33.2	5.9	509	6 US-09-925-065A-248642	Sequence 248642, A
44	33.2	5.9	509	6 US-09-925-065A-248643	Sequence 248643, A
45	33.2	5.9	514	10 US-10-301-480-328622	Sequence 328622, A

ALIGNMENTS

RESULT 1
US-11-121-086-61/c

Sequence 61, Application US/11121086
Publication No. US20050266459A1

GENERAL INFORMATION:

APPLICANT: NIELSEN, TIM S.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

FILE REFERENCE: 09138, 6000-00000

CURRENT FILING DATE: 2005-05-04

PRIOR APPLICATION NUMBER: 60/567,570

PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn version 3.3

SEQ ID NO 61

LENGTH: 169495

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: modified_base

LOCATION: (70072)..(70171)

OTHER INFORMATION: a, c, g, t, unknown or other

FEATURE:

NAME/KEY: modified_base

LOCATION: (139457)..(157244)

OTHER INFORMATION: a, c, g, t, unknown or other

US-11-121-086-61

Query Match 97.8% Score 551.4; DB 14; Length 169495;

Best Local Similarity 99.6%; Pred. No. 1.66-137; Indels 1; Gaps 1;

Matches 563; Conservative 0; Mismatches 1;

QY 1 GCAGCCCTGCCTCAGAAACGAGAGCGACGACACTCAGGTGACTCAACCCCATGTG 60

DB 51758 GCAGCCCTGCCTCAGAAACGAGAGCGACGACACTCAGGTGACTCAACCCCATGTG 51699

QY 61 GCTTGAGGTGAGGAGCCTCTCTGAGCGACGAGCGACGAGCGTCAAGTGGGTGACGCGCA 120

DB 51698 GCTTGAGGTGAGGAGCCTCTCTGAGCGACGAGCGACGAGCGTCAAGTGGGTGACGCGCA 51639

QY 121 GGGGCTTGGCCATGATGGGCAAGGGGCTGACATACGCTTACTCACTGACAAATCGAGTCC 180
DB 51638 GGGGCTTGGCCATGATGGGCAAGGGGCTGACATACGCTTACTCACTGACAAATCGAGTCC 51579
QY 181 CTGGTGGCCAGCTCTTGGAAAGTCTGGAAGTGAAGCAATGTTTCCCATTAAGAAAGTGTG 240
DB 51578 CTGGTGGCCAGCTCTTGGAAAGTCTGGAAGTGAAGCAATGTTTCCCATTAAGAAAGTGTG 51519
QY 241 GCGGCGCATATGCCCCCAAGTTGACATCACTGCTCTTGGCAAGGTTGGGGCTTCCAGTC 300
DB 51518 GCGGCGCATATGCCCCCAAGTTGACATCACTGCTCTTGGCAAGGTTGGGGCTTCCAGTC 51459
QY 301 ACAGGGTCCCATTCACGTACCAAGCCAGGTGCTGCAAGAGTCCCTGCAATCATGA 360
DB 51458 ACAGGGTCCCATTCACGTACCAAGCCAGGTGCTGCAAGAGTCCCTGCAATCATGA 51399
QY 361 CCAGGGAGGCTTGGAAACCAATCTGAAGGGCATGGCTTTGATTATTAAGAGAGGTGG 420
DB 51398 CCAGGGAGGCTTGGAAACCAATCTGAAGGGCATGGCTTTGATTATTAAGAGAGGTGG 51339
QY 421 GGCTGGGCTGGGCAAGGCCACAGTCTGATGACAGCAAGGCAAGGCAAGGCTGCTCCC 480
DB 51338 GGCTGGGCTGGGCAAGGCCACAGTCTGATGACAGCAAGGCAAGGCAAGGCTGCTCCC 51279
QY 481 A-GCACTGCCCCGCGCTTCTGCAATGCACTCTCTTGGCCACCTGGAACAGCTGTAGA 539
DB 51278 AGGCACTGCCCCGCGCTTCTGCAATGCACTCTCTTGGCCACCTGGAACAGCTGTAGA 51219
QY 540 GAGGCACTGCGCTCTTTCGGAATTC 564
DB 51218 GAGGCACTGCGCTCTTTCGGAATTC 51194

RESULT 2

US-09-925-065A-734471
; Sequence 734471, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 734471
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-734471

Query Match 7.1%; Score 39.8; DB 6; Length 541;
Best Local Similarity 64.8%; Pred. No. 0.7;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 4 GCGCTGCTCAGAAAACAGAGGACGACACCTACGCGTGACTACCCCATGTGCT 63
DB 442 GCGCTGCTCAGAAAACAGAGGAGCTGGGACCCCTCGTGGGCTCACCTCATTTTGAA 501
QY 64 GGAAGTGAAGGAGCTCTCTGAGGAGGAGGCA 94
DB 502 TGATGAGAGGTGTTCCCAAGGCTGGGCA 532

RESULT 3
US-11-121-086-20/c

; Sequence 20, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 162289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-20

Query Match 7.0%; Score 39.6; DB 14; Length 162289;
Best Local Similarity 46.9%; Pred. No. 2.1;
Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 3 AGCCCTGCTCAGAAAACAGAGGACGACACTCAAGGTGACTCAACCCCATGTGGC 62
DB 145401 AGCTAGCAGCACTGCTGGGAGGACCACTCTGCTGTGATGGGATCAGAGAG 145342
QY 63 TGAAGTGAAGGAGCCTCTGAGGAGGCGCAGGACCCGTACGTGGGTGACGCGAAG 122
DB 145341 CCTAGTCAAGGAGGAGGAGCTGAGGCGCCTCAAGAGAGTGGGTGGGCAATG 145282
QY 123 GGTCTTCCATGATGGGACAGGGCTGCTACACCTTACTATGATGACAAATGATCCT 182
DB 145281 GTCATGGGAGGTGGGACAGGTCAAGGCTGTGAGCACTTCCCAAGGCCAATGACAAAT 145222
QY 183 GGTGCGAGCTCTGGAAGTCTGGAAGTGAAGCAATGTTCCCATTAAGAAAGTGTGGC 242
DB 145221 GAGCAACCTTTGGGACTGACAGACATGTGGGGGCTTGGAGCCAGAGGAACTTAAGTGTG 145162
QY 243 CGCCATGCCCCCAGCGTTGC 264
DB 145161 ACACAAACACCCAGAGCTCTGC 145140

RESULT 4

US-10-750-185-32987
; Sequence 32987, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32987
; LENGTH: 712
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-32987

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 4.1;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 307 TCCATCCAGTACCAAGCCAGGTCCTGACAGAGTCCCTGACATGAACCAAG 366
DB 283 TCCATCCAGTACCAAGTCCATATGAGAAATTAATGAAGATGAAGATGCTG 342
QY 367 GAGGCTGGGAAACCATCTGAGAGGCGATGCTTTGATTTGTGAGAGGCGGCTG 426
DB 343 AAAAGTGTAAAGCATATTAAGCATTAATGCTTATGTAAGAAAGAGTGGCTCT 402
QY 427 GCTGGGCAAGCCAGGCTGAGTCAAGAGCAGAGCAAGAGTGGTCCCAAGCT 486
DB 403 TGCTTACTCCCTTACCTGAAATAGGCGAAGCGAGTATCCAGGATCTCTCAAGCT 462
QY 487 GCGGCGGCTCTGCGATGCAAGTCTCTGCGCCAGCTGAAACAGCTG 535
DB 463 TGCTTACTCTCTGCTTCTCATATTCATAGAGAGCTCGAAGAGACTG 511

RESULT 5

US-10-750-623-32987
; Sequence 32987, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISB, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32987
; LENGTH: 712
; TYPE: DNA
; ORGANISM: Bovine 19866880739882
US-10-750-623-32987

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 4.1;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 307 TCCATCCAGTACCAAGCCAGGTCCTGACAGAGTCCCTGACATGAACCAAG 366
DB 283 TCCATCCAGTACCAAGTCCATATGAGAAATTAATGAAGATGAAGATGCTG 342
QY 367 GAGGCTGGGAAACCATCTGAGAGGCGATGCTTTGATTTGTGAGAGGCGGCTG 426
DB 343 AAAAGTGTAAAGCATATTAAGCATTAATGCTTATGTAAGAAAGAGTGGCTCT 402
QY 427 GCTGGGCAAGCCAGGCTGAGTCAAGAGCAGAGCAAGAGTGGTCCCAAGCT 486
DB 403 TGCTTACTCCCTTACCTGAAATAGGCGAAGCGAGTATCCAGGATCTCTCAAGCT 462
QY 487 GCGGCGGCTCTGCGATGCAAGTCTCTGCGCCAGCTGAAACAGCTG 535
DB 463 TGCTTACTCTCTGCTTCTCATATTCATAGAGAGCTCGAAGAGACTG 511

RESULT 6
US-10-995-561-63329/c
; Sequence 63329, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63329
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-63329

Query Match 6.3%; Score 35.8; DB 8; Length 201;
Best Local Similarity 56.6%; Pred. No. 6.9;
Matches 64; Conservative 1; Mismatches 48; Indels 0; Gaps 0;

QY 45 GACTACCCCATGCTGAGAGTGAAGGAGGCTCTGAGGAGGAGGAGGAGGAGGCT 104
DB 138 GACTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79
QY 105 CAGGTGAGTGAAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 157
DB 78 TGGGTGGGAGTGGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 26

RESULT 7

US-10-995-561-69610
; Sequence 69610, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69610
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-69610

Query Match 6.3%; Score 35.8; DB 8; Length 201;
Best Local Similarity 56.6%; Pred. No. 6.9;
Matches 64; Conservative 1; Mismatches 48; Indels 0; Gaps 0;

QY 45 GACTACCCCATGCTGAGAGTGAAGGAGGCTCTGAGGAGGAGGAGGAGGAGGCT 104
DB 64 GACTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 105 CAGGTGAGTGAAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 157
DB 124 TGGGTGGGAGTGGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 176

RESULT 8
US-10-995-561-13412
; Sequence 13412, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561


```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: fastseq for windows version 4.0
; SEQ ID NO 707005
; LENGTH: 1741
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-707005
```

```

Query Match      6.2%; Score 35; DB 10; Length 1741;
Best Local Similarity 51.6%; Pred. No. 16;
Matches 80; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```

Qy 72 GGGAGCTCTGGAGCGAGCGGCGAGCGCTGAGCGGAGCGGCTTGGCC 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 GGGATGATGCTGACACAGGAGCGCTGAGGCGCTGGAGAGGCAAGGAGGCTTCC 131

Qy 132 ATGGTGGGACAGGGGCTGCAATGAGCTTACTGAGCAATTCGAGTCCCTGGTCCAGC 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 AGAGCTTCAAGGGGAGCATGCTGCTGACATCTTGAGCTGACAGCATCTGGCTTCAG 71

Qy 192 CTCTGGAAGTCTGGAAGTGAAGCAATGTTCCATT 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 GTATGACAGAGTGCATTTCCGTGAAGTGGCCCACT 36
```

```

RESULT 13
US-10-750-185-42100/c
; Sequence 42100, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42100
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Bovine 19866880864778
US-10-750-185-42100
```

```

Query Match      6.2%; Score 34.8; DB 8; Length 835;
Best Local Similarity 50.6%; Pred. No. 16;
```

```

Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
```

```

Qy 164 CAGTGAACATCGATCCTCTGTGCGAGCCTCTGGAAGTGTGAAGTGAATGTTTCCC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 CAGCAACGAGCGGCTCTGATGTCTAGTATGGCAGCAGCAGTGGAGTGAACCTG 479

Qy 224 ATTAAGAAAGTGTGTGGCCGCGCCATGCCCCCAACGTTGACACATCACTGCTTTGAG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 GGCAGGGAAGGTGTGCTGCTGGATGACTCAAGCTCAGCCCAATCACTCACTGCAA 419
```

```

Qy 284 GATTGGGGCTTCCAGTCACAGGTCCTCCATCCAGCTACAGCCAGG 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 GGTGTGTGACCCAGGCCCAACTTCAATCACTCTCTCCATCAGG 373
```

```

RESULT 14
US-10-750-623-42100/c
; Sequence 42100, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42100
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Bovine 19866880864778
US-10-750-623-42100
```

```

Query Match      6.2%; Score 34.8; DB 8; Length 835;
Best Local Similarity 50.6%; Pred. No. 16;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
```

```

Qy 164 CAGTGAACATCGATCCTCTGTGCGAGCCTCTGGAAGTGTGAAGTGAATGTTTCCC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 CAGCAACGAGCGGCTCTGATGTCTAGTATGGCAGCAGCAGTGGAGTGAACCTG 479

Qy 224 ATTAAGAAAGTGTGTGGCCGCGCCATGCCCCCAACGTTGACACATCACTGCTTTGAG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 GGCAGGGAAGGTGTGCTGCTGGATGACTCAAGCTCAGGCCCAATCACTCACTGCAA 419

Qy 284 GATTGGGGCTTCCAGTCACAGGTCCTCCATCCAGCTACAGCCAGG 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 GGTGTGTGACCCAGGCCCAACTTCAATCACTCTCTCCATCAGG 373
```

```

RESULT 15
US-09-925-065A-57042/c
; Sequence 57042, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: fastseq for windows version 4.0
; SEQ ID NO 57042
; LENGTH: 1172
```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-57042

Query Match 6.2%; Score 34.8; DB 6; Length 1172;
Best Local Similarity 58.8%; Pred. No. 17;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Oy 25 GGACCGACACACTCAGGTGACTCAGCCCATGTGCTGAGGTGAGGGAGCCTCTGA 84
|||
Db 599 GGGGGCAGCCCTGGCATGTGACTTTGGCAGATGAGAGAGACACAGAGCAGCGAGA 540
Oy 85 GGCAGGGCCAGGGCAGCCCTCAGGTGGGTGACGGCAGGGGTC 126
|||
Db 539 GGAAGGAGAGCGCGACCGCACGCGAGGTGACAGGTGACATC 498

Search completed: March 25, 2006, 18:31:07
Job time : 441.5 secB

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:16:12 ; Search time 681.5 Seconds
(without alignments)
6843.623 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtcgctcttcggacttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications NA Main:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	563	99.8	564	US-10-798-652-1	Sequence 1, Appl1
2	562	99.6	700	US-10-027-632-142127	Sequence 142127,
3	562	99.6	700	US-10-027-632-142127	Sequence 142127,
4	552.6	98.0	855	US-10-027-632-142126	Sequence 142126,
5	552.6	98.0	855	US-10-027-632-142126	Sequence 142126,
6	551.4	97.8	29340	US-10-322-281-642	Sequence 642, App
7	39.8	7.1	541	US-09-925-065A-734471	Sequence 734471,
8	38.2	6.8	2777	US-10-108-260A-578	Sequence 578, App
9	37.8	6.7	594	US-10-123-155-10	Sequence 10, Appl
10	37.8	6.7	594	US-10-146-731-10	Sequence 10, Appl
11	37.8	6.7	594	US-10-140-472-10	Sequence 10, Appl
12	37.8	6.7	594	US-10-141-756-10	Sequence 10, Appl
13	37.8	6.7	594	US-10-142-885-10	Sequence 10, Appl
14	37.8	6.7	594	US-10-158-790-10	Sequence 10, Appl
15	37.8	6.7	594	US-10-137-871-10	Sequence 10, Appl
16	37.8	6.7	594	US-10-140-923-10	Sequence 10, Appl
17	37.8	6.7	594	US-10-141-756-10	Sequence 10, Appl
18	37.8	6.7	594	US-10-141-759-10	Sequence 10, Appl
19	37.8	6.7	594	US-10-140-805-10	Sequence 10, Appl
20	37.8	6.7	594	US-10-140-864-10	Sequence 10, Appl
21	37.2	6.6	712	US-09-823-245A-162	Sequence 162, App
22	37.2	6.6	1957	US-09-764-847-1311	Sequence 1311, Ap
23	37.2	6.6	1957	US-09-764-891-10141	Sequence 10141, A

24	37.2	6.6	1957	5	US-10-092-154-1311	Sequence 1311, Ap
25	36.2	6.4	138203	9	US-10-819-386A-1	Sequence 1, Appl1
26	35.8	6.3	201	7	US-10-741-601-20932	Sequence 20932, A
27	35.8	6.3	201	7	US-10-741-600-60160	Sequence 60160, A
28	35.8	6.3	86361	7	US-10-741-601-5702	Sequence 5702, Ap
29	35.8	6.3	86361	8	US-10-741-600-17803	Sequence 17803, A
30	35.6	6.3	568	8	US-10-357-930-52013	Sequence 52013, A
31	35.6	6.2	550	8	US-09-918-995-12287	Sequence 12287, A
32	35	6.2	1338	6	US-09-843-905A-11	Sequence 11, Appl
33	35	6.2	1338	6	US-10-317-250-11	Sequence 11, Appl
34	35	6.2	1338	7	US-10-258-703-11	Sequence 11, Appl
35	35	6.2	1741	5	US-10-027-632-258790	Sequence 258790,
36	35	6.2	1741	6	US-10-027-632-258790	Sequence 258790,
37	35	6.2	1797	3	US-09-764-847-1312	Sequence 1312, Ap
38	35	6.2	1797	3	US-09-764-891-10143	Sequence 10143, A
39	35	6.2	1797	5	US-10-092-154-1312	Sequence 1312, Ap
40	34.8	6.2	1792	8	US-10-767-795-3312	Sequence 3312, Ap
41	34.8	6.2	1172	4	US-09-925-065A-57042	Sequence 57042, A
42	34.8	6.2	1198	4	US-09-925-065A-42328	Sequence 42328, A
43	34.8	6.2	1198	4	US-09-925-065A-42329	Sequence 42329, A
44	34.8	6.2	1533	4	US-09-925-065A-685510	Sequence 685510,
45	34.8	6.2	1533	4	US-09-925-065A-685511	Sequence 685511,

ALIGNMENTS

RESULT 1
US-10-798-652-1
; Sequence 1, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICATION: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-PO336US01
; CURRENT FILING DATE: 2004-03-11
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/455,698
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: n is cytosine or thymine
US-10-798-652-1
Query Match 99.8%; Score 563; DB 8; Length 564;
Beet Local Similarity 99.8%; Pred. No. 3e-164;
Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCAGCCCTGCTCGAGAAACGAGAGCGACACTCAGGTGACTCACCCCATYTG 60
DB 1 GCAGCCCTGCTCGAGAAACGAGAGCGACACTCAGGTGACTCACCCCATYTG 60
QY 61 GCTGAGAGTGGAGGAGCTCTTGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGA 120
DB 61 GCTGAGAGTGGAGGAGCTCTTGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGA 120
QY 121 GGGGCTTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 121 GGGGCTTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 181 CTGGTCCAGAGCTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTG 240
DB 181 CTGGTCCAGAGCTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTG 240
QY 241 GCGGCGATGCGCCCGCCCAAGCTGACACTGCTGCTTGGAGGATTGGGCTTCAAGTC 300

```
Db 241 GCCGCGCAATGCCCGCCCAAGCTTGACACATCTGCTTTGCAAGGGTTGGGGCTTCCAGTC 300
Qy 301 ACAGGGTCCCATCAGATACAGGACCCAGAGGAGGCTGAGAGAGGCTCCGCAATGATGAA 360
Db 301 ACAGGGTCCCATCAGATACAGGACCCAGAGGAGGCTGAGAGAGGCTCCGCAATGATGAA 360
Qy 361 CCAGGAGAGGCTTGGGAAACCAATCTGAAGGGCATGGCTTTGATTAGTAGAGGGGTG 420
Db 361 CCAGGAGAGGCTTGGGAAACCAATCTGAAGGGCATGGCTTTGATTAGTAGAGGGGTG 420
Qy 421 GGGTGGGCTGGGCAAGGCCACAGGCTCTGATGAGGCCAGAGGCGAGAAAGCTGGTCCCC 480
Db 421 GGGTGGGCTGGGCAAGGCCACAGGCTCTGATGAGGCCAGAGGCGAGAAAGCTGGTCCCC 480
Qy 481 AGCACTGCCCGCGGCTCTGCGATGAGTCTCTGAGCCACCTGAGAAACAGGCTGTAGAG 540
Db 481 AGCACTGCCCGCGGCTCTGCGATGAGTCTCTGAGCCACCTGAGAAACAGGCTGTAGAG 540
Qy 541 AGGCACTGGCGCTCTTTCGGACTTC 564
Db 541 AGGCACTGGCGCTCTTTCGGACTTC 564
```

```
RESULT 2
US-10-027-632-142127
; Sequence 142127, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142127
```

Query Match 99.6%; Score 562; DB 5; Length 700;

Best Local Similarity 99.6%; Pred. No. 6,4e-164;

Matches 562; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 60
Db 107 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 166
Qy 61 GCTGAGGTGAGGAGGCTCTTCTGAGGAGGCGCAAGGCGTCAAGTGGGTGACGGCA 120
Db 167 GCTGAGGTGAGGAGGCTCTTCTGAGGAGGCGCAAGGCGTCAAGTGGGTGACGGCA 226
Qy 121 GGGGTCTTGCATGTTGGGCAAGGGCTGGCATACAGCTTACATGAGCAATCGAGTCC 180
Db 227 GGGGTCTTGCATGTTGGGCAAGGGCTGGCATACAGCTTACATGAGCAATCGAGTCC 286
```

```
Qy 181 CTGGTCCAGGCTCTTGAGAAAGTGTGAGCAATGTTTCCATTAAAGAAAGTGTG 240
Db 287 CTGGTCCAGGCTCTTGAGAAAGTGTGAGCAATGTTTCCATTAAAGAAAGTGTG 346
Qy 241 GCCGGCCATGGCCCCCAAGTGTGCACTCACTGCTTTGCAAGGGTTGGGGCTTCCAGTC 300
Db 347 GCCGGCCATGGCCCCCAAGTGTGCACTCACTGCTTTGCAAGGGTTGGGGCTTCCAGTC 406
Qy 301 ACAGGGTCCCATCAGATACAGGACCCAGAGGAGGCTGAGAGAGGCTCCGCAATGATGAA 360
Db 407 ACAGGGTCCCATCAGATACAGGACCCAGAGGAGGCTGAGAGAGGCTCCGCAATGATGAA 466
Qy 361 CCAGGAGAGGCTTGGGAAACCAATCTGAAGGGCATGGCTTTGATTAGTAGAGGGGTG 420
Db 467 CCAGGAGAGGCTTGGGAAACCAATCTGAAGGGCATGGCTTTGATTAGTAGAGGGGTG 526
Qy 421 GGGTGGGCTGGGCAAGGCCACAGGCTCTGATGAGGCCAGAGGCGAGAAAGCTGGTCCCC 480
Db 527 GGGTGGGCTGGGCAAGGCCACAGGCTCTGATGAGGCCAGAGGCGAGAAAGCTGGTCCCC 586
Qy 481 AGCACTGCCCGCGGCTCTGCGATGAGTCTCTGAGCCACCTGAGAAACAGGCTGTAGAG 540
Db 587 AGCACTGCCCGCGGCTCTGCGATGAGTCTCTGAGCCACCTGAGAAACAGGCTGTAGAG 646
Qy 541 AGGCACTGGCGCTCTTTCGGACTTC 564
Db 647 AGGCACTGGCGCTCTTTCGGACTTC 670
```

```
RESULT 3
US-10-027-632-142127
; Sequence 142127, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142127
```

Query Match 99.6%; Score 562; DB 6; Length 700;

Best Local Similarity 99.6%; Pred. No. 6,4e-164;

Matches 562; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 60
Db 107 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 166
Qy 61 GCTGAGGTGAGGAGGCTCTTCTGAGGAGGCGCAAGGCGTCAAGTGGGTGACGGCA 120
Db 167 GCTGAGGTGAGGAGGCTCTTCTGAGGAGGCGCAAGGCGTCAAGTGGGTGACGGCA 226
```

Qy	121	GGGGCTTTCAGTAGGGGACAGGGGCTGCATACAGCTTACTCAGTGAACATCGAGTCC	180
Db	227	GGGGCTTTCAGTAGGGGACAGGGGCTGCATACAGCTTACTCAGTGAACATCGAGTCC	286
Qy	181	CTGTGTCCAGCCTCTTGGAAGTCTGGAAGTGAAGCAATGTTTCCCATTAAGGAAAGTGTG	240
Db	287	CTGTGTCCAGCCTCTTGGAAGTCTGGAAGTGAAGCAATGTTTCCCATTAAGGAAAGTGTG	346
Qy	241	GGCGGCAATGGCCCCCAAGTTGCACATCATGCCCTTTGCAGGGTTGGGGCTTCCAGTC	300
Db	347	GCCGGCATATGCCCCCAAGCTTGCACATCATGCTTTTGCAGGGTTGGGGCTTCCAGTC	406
Qy	301	ACAGGGTCCCATCCAGTACCAAGCCAGGTGTGCTGCAGAAAGTCCCTCCAGTCATGAAA	360
Db	407	ACAGGGTCCCATCCAGTACCAAGCCAGGTGTGCTGCAGAAAGTCCCTCCAGTCATGAAA	466
Qy	361	CCAAAGGAGAGCTTTGGGAAACCACTTTGAAAGGCATGGCTTTGATTAGTAGAGGGTGG	420
Db	467	CCAAAGGAGAGCTTTGGGAAACCACTTTGAAAGGCATGGCTTTGATTAGTAGAGGGTGG	526
Qy	421	GGCTGGGCTGTGGCAAGGCCACAGGCTGTAGTCAGAGCCAGAGGACAGAAAGTGTGCCCC	480
Db	527	GGCTGGGCTGTGGCAAGGCCACAGGCTGTAGTCAGAGCCAGAGGACAGAAAGTGTGCCCC	586
Qy	481	AGCACTGCCGCGCCTCTTGCGATGTGACGTCTTGTGGCAACTGTGAAGCAGCCTGTAGAG	540
Db	587	AGCACTGCCGCGCCTCTTGCGATGTGACGTCTTGTGGCAACTGTGAAGCAGCCTGTAGAG	646
Qy	541	AGGCAGTGGCGTCTTTGGAGCTTC	564
Db	647	AGGCAGTGGCGTCTTTGGAGCTTC	670

RESULT 4
US-10-027-632-142126
; Sequence 142126, Application US/10027632
; Publication No. US20020198371A1
; General Information

```

? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? TITLE OF INVENTION: Polymorphisms in the Human Genome
? FILE REFERENCE: 108827,129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218,006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 142126
? LENGTH: 855
? TYPE: DNA
? ORGANISM: Human
? US-10-027-632-142126

```

Query Match	98.0%	Score 552.6	DB 5	Length 855
Best Local Similarity	99.6%	Pred. No. 5.4e-161		
Matches 563	1	Mismatches 0	Indels 1	Gaps 1
QY	1	GCAGCGCTCCTCAGAAACAGAGAGGACGACATCCAGCGTACCTACCCCCATTTG	60	

Db	108	GGAGGCTTCGCTCAGAAAAACAAAGGACGACACACTACGGTGACTACCCCATGTG	167
Qy	61	GCTGAGAGTGAAGAGAGCTCTCTGAGGACAGGAGCCAGGAGCCGCTCAGGTGGGTGAAGGCA	120
Db	168	GCTGAGAGGGAAGAGAGCTCTCTGAGGACAGGAGCCAGGAGCCGCTCAGGTGGGTGAAGGCA	227
Qy	121	GAGGCTTGGCCATGATGGGCA CAGGGGCTGCATACAGTTACTCAATGACAAATCGAGTCC	180
Db	228	GAGGCTTGGCCATGATGGGCA CAGGGGCTGCATACAGTTACTCAATGACAAATCGAGTCC	287
Qy	181	CTGTGTCCAGCCTCTGGAAGTCTGGAAGTGAAGCAATGTGTTCCATTMAAGAAAGTGTG	240
Db	288	CTGTGTCCAGCCTCTGGAAGTCTGGAAGTGAAGCAATGTGTTCCATTMAAGAAAGTGTG	347
Qy	241	GCCGCGCATGCCCCCAACGTTGCACTACTGCTTTGCAAGGGTTGGGGCTTCCAGTC	300
Db	348	GCCGCGCATGCCCCCAACGTTGCACTACTGCTTTGCAAGGGTTGGGGCTTCCAGTC	407
Qy	301	ACAGGGTCCCATCCACGTAAACGAGCCAGGTGCTCAGAAAGTTCCTGCAGTCAGAAA	360
Db	408	ACAGGGTCCCATCCACGTAAACGAGCCAGGTGCTCAGAAAGTTCCTGCAGTCAGAAA	467
Qy	361	CCAAAGGAGAGCTTGGGAAACCATCTGAAAGGAGCATGTGATTTAAGTAGAGGGTGG	420
Db	468	CCAAAGGAGAGCTTGGGAAACCATCTGAAAGGAGCATGTGATTTAAGTAGAGGGTGG	527
Qy	421	GAGTGGAGCTGGGCAAGGACCAACGAGTCTGAGTCAGAGCCAGAGGCAAGAAAGCTGTCCCC	480
Db	528	GAGTGGAGCTGGGCAAGGACCAACGAGTCTGAGTCAGAGCCAGAGGCAAGAAAGCTGTCCCC	587
Qy	481	A-GCACTGCCCCCGCCCTCTCTGAGATGCAATCTCTGGGCACCTGAGAAACAGCTGTAGA	539
Db	588	AGGCACTGCCCCCGCCCTCTCTGAGATGCAATCTCTCTGGGCACCTGAGAAACAGCTGTAGA	647
Qy	540	GAGGCAAGTGGGCTCTTTGCACTTC	564
Db	648	GAGGCAAGTGGGCTCTTTGCACTTC	672

RESULT 5
US-10-027-632-142126
US-10-027-632-142126
Sequence 142126, Application US/10027632
Publication No. US20030204075A
Publication Date 02/27/2003

```

; TITLE: Wang, David G.
; TYPE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 109827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/199,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142126
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-142126

```

Query Match 98.0%; Score 552.6; DB 6; Length 855;

Db 502 TGATGAGAGTGTCTCCCAAGGCTGGGCCA 532

RESULT 8

US-10-108-260A-578/c
; Sequence 578, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 578
; LENGTH: 2777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-578

Query Match 6.8%; Score 38.2; DB 6; Length 2777;
Best Local Similarity 48.8%; Pred. No. 0.25;

Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 79 TCCTGAGCAGGGCCAGGGCCGCTCAGGTGAGCGGCGGCTTGGCATGTGG 138
Db 345 TCCGCCCCAGAGGTGTGTCACAGTCTGGGGGGTCACTGGCGCAGCCAGGCTGGGGGG 286
Qy 139 GCACAGGGGCTGCATACAGCTTACTCAGTGAATGATGATCCCTGTGTCAGCTCTGGA 198
Db 285 CCACCGGAGCCCCACCTGCGGTGTCCACCCGAGCATCGCTTGGGCCCGGCACTCAG 226
Qy 199 AGTCTGAGAGTGAAGCAATGTTTCCCATTTAGGAAAGTGTGTGCGGCCCATGCCCCCA 258
Db 225 AGTGAAGCAGAGGCACTTGTCTCTCTTGTATGGCACCCAGGCGCCAGGAGCCGATGAC 166
Qy 259 CGTTCACACCTCAGTCCCTTGGCAGAGGTGG 289
Db 165 CATGACAAAGCGCCGCGTGTATGATGAGCAGG 135

RESULT 9

US-10-123-155-10
; Sequence 10, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123, 155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10

; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-10

Query Match 6.7%; Score 37.8; DB 5; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26;

Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

Qy 23 AAGACGACGACACTCAGGTGACTACCCCCATGTGGCTGAGAGTGAAGGACCTTCT 82
Db 56 VALAESPVNMWPMQGDNNMIDRPVRAHMDHPDYPPLITTSPEQBEDERKCYNER 115
Qy 83 GAGGACGAGCCAGGACCGCTCAGGTGAGCGGCGGGGTCTTGCATGTGGGAC 142
Db 116 YRGVQNDPAGISEBQCIYIYIDELVGLRPSBEDKKLAEKKASIGTYEDSTVAEV 175
Qy 143 AGGGCTGCATACAGCTTACAGTGAATGATCCCTGTGTCAGCTCTGGAAGTC 202
Db 176 EKAAREPPEBSAABEBSNDEVIYDIDVEDVDLNOQVADLNQATTYGWAQDF 235
Qy 203 TGAAGTGAACAATGTTTCCCATTAAGAAAGTGTGCGCGCCATGCCCCCAAGTT 262
Db 236 VRMRKQDEBAEALKHAVALBEERAMYSGRSRQRREFRERKRLGRKISPPSYARDSP 295
Qy 263 GCACACTCAGTCCCTTTCAGAGGTGGGCTTCCAGTACAGGGTCCCATCAGCTACCA 322
Db 296 TYDPKPSPSRSSSSRSRSRSPTPGREKITFTTSPGSEBEAAMAAAAAASGVTTGK 355
Qy 323 GCCCAGGTGCTGACGAAGTCCCTCGCATGATGAACCAAGGAGGCTTGGGAAACA 382
Db 356 PPAPQPGAPAPGNASARRSSSSSSSSSRSSSSSSSSSSRRSGGYYRSGHA 415
Qy 383 CATCGAAGGACATGCTTTATTTAGTGAAGGTGGGCTGGCTGGCGCAAGCCAC 442
Db 416 RSRSRWSRSRSTRSRSRSRSGGSRGHRYSRSPARRGYGPARRSRSRSHS 475
Qy 443 AGTCTGAGTCAAGCAGAGCAGGAAGTCCGACGACGCTCCGACGCTCCGCTTGG 502
Db 476 GDRYRRGGRLHHSSSRSSWSLSPSRKSLTRSHSPSPQSRSRSSQSPSPS 535
Qy 503 ATGCAGT 509
Db 536 PAREKLT 542

RESULT 10

US-10-146-731-10
; Sequence 10, Application US/10146731
; Publication No. US20030128692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146, 731
; CURRENT FILING DATE: 2002-05-15

/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-146-731-10

Query Match
Best Local Similarity 7.0%; Score 37.8; DB 6; Length 594;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

23 AAGGACGACGACACTCAGGAGTCAAGCCCATGAGTGGTGGAGGTGAGGAGGCTCCT 82
56 VALAAESPVMWPMQDNTNMIDRFPVRAHLDHPYTPPLTTTSPQESDERKCNTER 115
83 GAGGACGAGGCGCAGGAGCGGTGAGTGGTGGAGGAGGCTTGGTGGAGGAGC 142
116 YRGLVQNDPAGISEEQCLYQIYIDELYGGLQRPSEDEKKLAERKASIGTYEDSTVAEV 175
143 AAGGGCTGATACAGCTTACTGACATGACATGAGTCCGTGGTGGCAGGCTGTGGAATC 202
176 EKAAEKPEEESAAEESNSDEDEVIPDIDVEVDVDELNOBQVADLNKQATTYGMADGDF 235
203 TGGAGTGACGACATGTTTCCATTAAAGAAAGTGTGGCGGCGCATGCCCCCAACGTT 262
236 VMRLKDXEAEALIKAAKLEBEKAMYSGRRSRQRFREKRLRGRKISPSYARBDP 295
263 GCACACTACTGCTTTCAGAGGTTGGGCTTCACATCAGAGGTCCTCCATCCAGTACCA 322
296 TYDPYKSPSESSSRSRSPPTPGRBKITFTSPGSDDEBAAAAAAGVTTGK 355
323 GCCCAGTGTGTCAGAGGTCCTGTCAGTCAAGTCAAGAGGAGGAGGCTGGGAAACCA 382
356 PPAPPOGPGAPARNAASARRSSSSSSASRTSSSRSSSRSGGTYRSGRA 415
383 CATCTGAAGGCGATGCTTGTATTAGTAGAGAGGTGGGCTGGGCGGAGGCGCAC 442
416 RSRSRWSRSRSRSRSRSGRRHSGGSRDGHYRSPARRGGYTPRRSRSRSHS 475
443 AGGTCTGAGTCAGAGCGCAGAGGAGGAGTGTCCCGACACTGCCCGCCCTCTGGC 502
476 GDRYRGGRGRLRHSSSRSSWSLSPSRSLTTRSRSHSPSQSRSRSSQSPSPS 535
503 ATGCAGT 509
536 PARERLT 542

RESULT 11
US-10-140-472-10
/ Sequence 10, Application US/10140472
/ Publication No. US2003013888A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Matanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME

/ FILE REFERENCE: P3330R1C168
/ CURRENT APPLICATION NUMBER: US/10/140,472
/ CURRENT FILING DATE: 2002-05-06
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-472-10

Query Match
Best Local Similarity 7.0%; Score 37.8; DB 6; Length 594;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

23 AAGGACGACGACACTCAGGAGTCAAGCCCATGAGTGGTGGAGGTGAGGAGGCTCCT 82
56 VALAAESPVMWPMQDNTNMIDRFPVRAHLDHPYTPPLTTTSPQESDERKCNTER 115
83 GAGGACGAGGCGCAGGAGCGGTGAGTGGTGGAGGAGGCTTGGTGGAGGAGC 142
116 YRGLVQNDPAGISEEQCLYQIYIDELYGGLQRPSEDEKKLAERKASIGTYEDSTVAEV 175
143 AAGGGCTGATACAGCTTACTGACATGACATGAGTCCGTGGTGGCAGGCTGTGGAATC 202
176 EKAAEKPEEESAAEESNSDEDEVIPDIDVEVDVDELNOBQVADLNKQATTYGMADGDF 235
203 TGGAGTGACGACATGTTTCCATTAAAGAAAGTGTGGCGGCGCATGCCCCCAACGTT 262
236 VMRLKDXEAEALIKAAKLEBEKAMYSGRRSRQRFREKRLRGRKISPSYARBDP 295
263 GCACACTACTGCTTTCAGAGGTTGGGCTTCACATCAGAGGTCCTCCATCCAGTACCA 322
296 TYDPYKSPSESSSRSRSPPTPGRBKITFTSPGSDDEBAAAAAAGVTTGK 355
323 GCCCAGTGTGTCAGAGGTCCTGTCAGTCAAGTCAAGAGGAGGAGGCTGGGAAACCA 382
356 PPAPPOGPGAPARNAASARRSSSSSSASRTSSSRSSSRSGGTYRSGRA 415
383 CATCTGAAGGCGATGCTTGTATTAGTAGAGAGGTGGGCTGGGCGGAGGCGCAC 442
416 RSRSRWSRSRSRSRSGRRHSGGSRDGHYRSPARRGGYTPRRSRSRSHS 475
443 AGGTCTGAGTCAGAGCGCAGAGGAGGAGTGTCCCGACACTGCCCGCCCTCTGGC 502
476 GDRYRGGRGRLRHSSSRSSWSLSPSRSLTTRSRSHSPSQSRSRSSQSPSPS 535
503 ATGCAGT 509
536 PARERLT 542

RESULT 12
US-10-141-761-10
/ Sequence 10, Application US/10141761
/ Publication No. US20030148432A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Matanabe, Colin K
/ APPLICANT: Wood, William

```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR APPLICATION: 2002-05-08
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-141-761-10

```

```

Query Match      6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

```

```

QY 23 AAGACGACGACACTCAGGTGACTACCCCATGCTGAGGTGAGGAGGCTCT 82
DB 56 VALAESPVNMMPWQDNTNMIDRFVRAHLDHPYPLTTISPQESDERKCYNER 115
QY 83 GAGGACGAGGCGACGCGCTCAGGTGAGGTGAGGAGGCTCTGCGTGGGAC 142
DB 116 YRGLVQNPAGISBQCLYQIYIDELYGGLQRPSEDEKKALEKASIGTYEDSTVAEV 175
QY 143 AGGGGCTGCATACGCTTACTCAGTACGACATGACATGCTGCTGAGGCTCTGGAATC 202
DB 176 EKAKEPBEESAAEBESNDEVIPIDIVEDVDLQBVADLNKQATTYGMADDF 235
QY 203 TGAAGTGACATGATGTTCCATTAGGAAGTGTGCGCGCATGCGCCCAACGTT 262
DB 236 VRMLKDEBAEAIKHAVALKEEKAMYSGRSRQRREFREKRLRGKISPSYARRSP 295
QY 263 GCACACTCAGTCTGCTTGGAGGCTTGGGCTTCCAGTACAGGCTCCATCCAGTACCA 322
DB 296 TYDPKRSPPSSSSSRSSRSPPTGREGKITFTISFGSDEBAAAAAAASGVTGK 355
QY 323 GCCCAGGTGCTGCAGAAAGTCCCTCGCATGATGAACCAAGGAGGCTTGGAAACA 382
DB 356 PPAPQPGPAPGAPGNAARRSSSSSSASRTSSSSSSSSSRSGGYYRSGHNA 415
QY 383 CATCTGAAGGCGCATGCTTTGATTAGTGAAGGCTGGGCTGGGCTGGGCAAGCCAC 442
DB 416 RSRSSWSRSRSTRSRGRHSGGSRDHRYSRPARRGYGPARRSRSHS 475
QY 443 AGGTCTAGTCAAGACCGACGAGGAGGAGGCTGCCAGCATGCTCCGCGCTTGGC 502
DB 476 GDRYRGRGRLRHSSSRSSWSLSPSRSLTRSHSPSPQSRSSRSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

```

```

RESULT 13
US-10-142-885-10
; Sequence 10, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; PRIOR APPLICATION: 2002-05-10
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-142-885-10

```

```

Query Match      6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

```

```

QY 23 AAGACGACGACACTCAGGTGACTACCCCATGCTGAGGTGAGGAGGCTCT 82
DB 56 VALAESPVNMMPWQDNTNMIDRFVRAHLDHPYPLTTISPQESDERKCYNER 115
QY 83 GAGGACGAGGCGACGCGCTCAGGTGAGGTGAGGAGGCTCTGCGTGGGAC 142
DB 116 YRGLVQNPAGISBQCLYQIYIDELYGGLQRPSEDEKKALEKASIGTYEDSTVAEV 175
QY 143 AGGGGCTGCATACGCTTACTCAGTACGACATGACATGCTGCTGAGGCTCTGGAATC 202
DB 176 EKAKEPBEESAAEBESNDEVIPIDIVEDVDLQBVADLNKQATTYGMADDF 235
QY 203 TGAAGTGACATGATGTTCCATTAGGAAGTGTGCGCGCATGCGCCCAACGTT 262
DB 236 VRMLKDEBAEAIKHAVALKEEKAMYSGRSRQRREFREKRLRGKISPSYARRSP 295
QY 263 GCACACTCAGTCTGCTTGGAGGCTTGGGCTTCCAGTACAGGCTCCATCCAGTACCA 322
DB 296 TYDPKRSPPSSSSSRSSRSPPTGREGKITFTISFGSDEBAAAAAAASGVTGK 355
QY 323 GCCCAGGTGCTGCAGAAAGTCCCTCGCATGATGAACCAAGGAGGCTTGGAAACA 382
DB 356 PPAPQPGPAPGAPGNAARRSSSSSSASRTSSSSSSSSSRSGGYYRSGHNA 415
QY 383 CATCTGAAGGCGCATGCTTTGATTAGTGAAGGCTGGGCTGGGCTGGGCAAGCCAC 442
DB 416 RSRSSWSRSRSTRSRGRHSGGSRDHRYSRPARRGYGPARRSRSHS 475
QY 443 AGGTCTAGTCAAGACCGACGAGGAGGAGGCTGCCAGCATGCTCCGCGCTTGGC 502
DB 476 GDRYRGRGRLRHSSSRSSWSLSPSRSLTRSHSPSPQSRSSRSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

```

```

RESULT 14
US-10-158-790-10
; Sequence 10, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.

```

```

: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tunes, Daniel
: APPLICANT: Matanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zhenli
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3330R1C448
: CURRENT APPLICATION NUMBER: US/10/158,790
: PRIORITY FILING DATE: 2002-05-30
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 10
: LENGTH: 594
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-10-158-790-10

```

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0

Oy	23	AAAGACGACGACACTCACGGTACTCACCCCCATGCTGGAGTGAAGGACGCTCT	82
Db	56	VALAAESPVMWPGDITNMIDRPDPVRAHLHDIPDYPLTLTISPEDESEKCONYER	115
Oy	83	GAGGAGAGGCGCAGGGCAGCCGTCAGGTGGGTGAOCGACAGGGGCTCTTGCATGTTGGGAC	142
Db	116	YRGVQNPAGISIEBOCLYIYIDELVGGLORPSEDEKKLAEKTAISIGITYEDSTVAEV	175
Oy	143	AGGGGCTGCATCTACTAGTCAGTCACATCGAGCTCGGTGGCCAGCCTTGGACAGTC	202
Db	176	EKAAMEPEEBEBAAEBSNDSDEVYIPDIDVEDVDDELNOBVADLNMKATTYGMAADGDF	235
Oy	203	TGGAAGTGACGAAATGTTTCCATTAAAGAAATGTGTGGCGGACCATGCCCCAACGTT	262
Db	236	VRMLKDKBEBAVAIHAKALBEEKAMYSGRSRGRQREPRERKRLNRKIKSPSYARBDP	295
Oy	263	GCACACTCATGCTCTTTCGAGGGTGTGGGCTTCCAGTCACAGGCTCCCATCCAGTACCA	322
Db	296	TYDPYKRSPEBESSSSRSRSPTPGREREKITFTITSFGSGDEBAAAAAAAAAAGTTGK	355
Oy	323	GCCGAGGTGGCGTCAGAAAGTCCCTCCGCACTCATGAAACCAAGGGAGGCTTGGAAACA	382
Db	356	PPAPPOGPGAPAGRNASARRRSSSSSSSSSARSTSSRSRSSSRGGGYNSGRHA	415
Oy	383	CATCGAAGGCGATGGCTTTGATTGAATTAGTACAGAGGTGGCGCTCGCTGGCAAGGCACC	442
Db	416	RSRSISWSRSRSRSRRYRSRSRGRGRHSGGSRDGRHRYRSRPARRCGYPRRRRSRSRS	475
Oy	443	AGGTGTGAGTCAGACCCAGGCGAGGACGAGACGTGCTCCCGACGACTGCCCGCCTTGC	502
Db	476	GDRYRRGGGLRHHSSSRSSWSLSPSSRLSTRSSSPSPGSRSSRSRSGPSPS	535
Oy	503	ATGCAGT	509
Db	536	PAREKLT	542

US-10-137-871-10
US-10-137-871-10
Sequence 10, Application US/10137871
Publication No. US20030207350A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.

```

/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P330R1C153
/ CURRENT APPLICATION NUMBER: US/10/137,871
/ CURRENT FILING DATE: 2002-05-03
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
/ US-10-137-871-10

```

Query Match	6.7%	Score 37.8;	DB 6;	Length 594;
Best Local Similarity	7.0%;	Pred. NO. 0.26;		
Matches 34;	Conservative 176;	Mismatches 277;	Indels 0;	Gaps 0

QY	23	AAAGACGACGACACTACGGTAGTACACCCCCATGTGGCTGGAGGTAGGAGGAGCGCTCT	82
Db	56	VALAAESPVNMMPQMGDTNNMIDRFVRAHLDHIDPYRPLLTTTISPQESDERCKNYER	115
QY	83	GAGGACGGGCCGAGCGCAGCCGTACAGTGGGTGAACGGCAGGGGTCTTGCATGTGTACGCAC	142
Db	116	YRGVQNDNFAGISIEBQCLYQIYIDELVYGLQRPSEDEKKLAEKASIGYTEDSTVAEV	175
QY	143	AGGGGCTGCATACAGCTTACTCAGTGAACAATCGAGTCCCTGGTGCACAGCTCTGGAGTGC	202
Db	176	EKAAMEKPEEESAAAEESNSDEVIDIPIDIVENVVDLNGQVADLNMQATTTGAAADDF	235
QY	203	TGGAAGTAGCAATGTTTTCCATTAAAGAAAGTGTGTGCGCGGCAGTCCGCCCAACGTT	262
Db	236	VRMLRKDEBAEAIHAAKALBEEKAMVSGRRSRBQREFREKRLGRKISPSVARDSP	295
QY	263	GCACACTACCTGCTTTGCAGAGGTTGGGGCTTCCAGTACACAGGGTCCATCCAGTACCA	322
Db	296	TYDPKRSPPSESSSSRSRSPPTPGREKITTFTIFSFGSDSEBAAAAAASAVTTGK	355
QY	323	GCCCAAGTGGCTGCAGAGAGTCCCTCCGAGCATATAAACAAGAGGCTTGCGAAACCA	382
Db	356	PPAPPOGPPAPGRNVAASARRSSSSSSSSASRTSSSSSSSSRSRSGGTYSGRHA	415
QY	383	CATCGAAGAGCATGCTTTGATTATTAGTAGAGAGGTGGGCTGGCTGGCGCAAGCCACC	442
Db	416	RSRSMSWSRSRKRYSRSRSGRGRHSGGGSRDHAKRSRPARRGGTGPRRARSRSHS	475
QY	443	AGGTCTAGTACAGACGACGAGGACGAAGATGTGCTCCACGACCTGCCCGCGCTCTGCG	502
Db	476	GDRYVRGGRGLRHHSRSRSGSWLSIPRSRSLTRSRSHSPSQSRSRSRSSQSPSPS	535
QY	503	ATGCAGT 509	
Db	536	PARETILT 542	

Search completed: March 25, 2006, 18:14:35
Job time : 682.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 16:11:03 ; Search time 3009.5 Seconds
(without alignments)
10652.835 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaac.....agtgcgtcttcgacttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	560.8	99.4	10777	8	HS11013RP
2	551.4	97.8	150110	14	AC097722
3	551.4	97.8	150214	8	AF006345
4	43	7.6	125020	8	AF429315
5	41.4	7.3	125020	8	AF429315
6	40.2	7.1	320658	14	AC145527
7	40	7.1	153129	9	AC163309
8	40	7.1	230041	9	AC111047
9	39.8	7.1	103034	14	AC067978
10	39.8	7.1	111934	8	AF281074
11	39.8	7.1	116039	8	AC007362
12	39.8	7.0	123561	8	AC159874
13	39.6	7.0	119613	8	AC138582
14	39.6	7.0	142879	8	AC000029
15	39.6	7.0	190018	14	AC148588
16	38.8	6.9	184532	9	AC124345
17	38.8	6.9	203106	9	AC021667
18	38.8	6.9	215604	14	AC160979

c 19	38.6	6.8	184533	8	AC093592	AC093592 Homo sapi
c 20	38.6	6.8	192573	8	AC073344	AC073344 Homo sapi
c 21	38.4	6.8	110000	14	CT005255_5	Continuation (6 of
c 22	38.2	6.8	2777	6	AX833454	AX833454 Sequence
c 23	38.2	6.8	2777	8	AK095212	AK095212 Homo sapi
c 24	38.2	6.8	74871	14	AC137095	AC137095 Homo sapi
c 25	38.2	6.8	178778	8	AC132936	AC132936 Homo sapi
c 26	38	6.7	110000	14	LMFLCHR26_2	Continuation (3 of
c 27	38	6.7	121103	8	AP006623	AP006623 Homo sapi
c 28	38	6.7	205854	14	AC138186	AC138186 Homo sapi
c 29	37.8	6.7	113370	8	AC067815	AC067815 Homo sapi
c 30	37.8	6.7	134362	14	AC084853	AC084853 Homo sapi
c 31	37.8	6.7	138129	8	HS102D24	AL021391 Human DNA
c 32	37.8	6.7	186083	9	AC122522	AC122522 Mus muscu
c 33	37.8	6.7	200629	9	AC134463	AC134463 Mus muscu
c 34	37.6	6.7	114694	8	AL355500	AL355500 Human DNA
c 35	37.4	6.6	184866	9	AL611934	AL611934 Mouse DNA
c 36	37.4	6.6	187911	14	AC136867	AC136867 Rattus no
c 37	37.4	6.6	207735	9	AC130279	AC130279 Mus muscu
c 38	37.4	6.6	213590	14	AC154101	AC154101 Mus muscu
c 39	37.4	6.6	223357	14	AC111958	AC111958 Rattus no
c 40	37.4	6.6	240039	14	AC111926	AC111926 Rattus no
c 41	37.2	6.6	21670	8	AF508041	AF508041 Homo sapi
c 42	37.2	6.6	36628	8	AF037222	AF037222 Human DNA
c 43	37.2	6.6	294817	14	AC114626	AC114626 Mus muscu
c 44	37	6.6	1422	4	BT020874	BT020874 Bos tauru
c 45	37	6.6	2500	1	SS055940	SS055940 Streptomyce

ALIGNMENTS

RESULT 1	HS11013RP	10777 bp	DNA	linear	PRI 02-MAR-2000
LOCUS	HS11013RP				
DEFINITION	H.sapiens FGF-3 gene upstream flanking region.				
ACCESSION	Y12377				
VERSION	Y12377.1	GI:1934871			
KEYWORDS	Alu repeat; FGF-3 gene; int-2 gene; L1 repeat; MIR repeat.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Dienabi,S., Brisson,O., Galdemand,C. and Lavialle,C.			
AUTHORS		Sequence analysis of the transcription control region upstream of			
TITLE		the human FGF-3 gene			
JOURNAL		DNA Seq. 10 (4-5), 317-329 (1999)			
PUBMED		10727086			
REFERENCE		2 (bases 1 to 10777)			
AUTHORS		Brisson,O.			
TITLE		Direct Submission			
JOURNAL		Submitted (04-APR-1997) O. Brisson, Lab. de Genetique Oncologique,			
		CNRSURA1967, Institut Gustave-Roussy, 39 Rue Camille Desmoulins,			
		94805 Villejuif CEDEX, FRANCE			
		The BamHI site at the 3' end of this sequence (nt 10772)			
		corresponds to the BamHI site at the 5' end of the FGF-3/int-2 gene			
		sequence (nt 1; X1445), as ascertained by sequencing through this			
		site.			
		Related sequence: B04369.			
FEATURES		Location/Qualifiers			
source		1..10777			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
		/chromosome="11"			
		/map="q13"			
		/cissue_type="placenta"			
		/clone_id="lambda D3, Clontech cat#HL10675 lot 1221"			
		/germline			
		489..631			
		/rpt_family="MIR2"			
		repeat_region			

```
repeat_region 632..945
/rpt_family="Alu"
repeat_region 1084..1167
/rpt_family="L1MB7"
repeat_region 2605..2752
/rpt_family="Alu"
repeat_region 3224..3514
/rpt_family="Alu"
repeat_region 5800..6089
/rpt_family="Alu"
repeat_region 7003..7242
/rpt_family="MLT1B"
repeat_region 7417..7568
/rpt_family="L1MB3"
repeat_region 7715..7844
/rpt_family="MIR"
repeat_region
```

ORIGIN

```
Query Match 99.4%; Score 560.8; DB 8; Length 10777;
Best Local Similarity 99.6%; Pred. No. 2.2e-122;
Matches 562; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
1 GGAGCCCTGCTCAGAAAACAGAAAGACGACGACACCTGAGTACACCCCATGTC 60
|||||
4945 GGAGCCCTGCTCAGAAAACAGAAAGACGACGACACCTGAGTACACCCCATGTC 5004
61 GCTGAGGTGAGGAGGAGCTCTCTGAGGACAGGACGAGCCGTCAGTGGTGAACGCA 120
|||||
5005 GCTGAGGTGAGGAGGAGCTCTCTGAGGACAGGACGAGCCGTCAGTGGTGAACGCA 5064
121 GGGGCTTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
|||||
5065 GGGGCTTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5124
181 CTGGTGCCAGCCCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAGAAAGTGTG 240
|||||
5125 CTGGTGCCAGCCCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAGAAAGTGTG 5184
241 GCCGGCCATGCCCCCAAGTTCACACTGCTTTCAGAGGTTGGGGCTTCCAGTC 300
|||||
5185 GCCGGCCATGCCCCCAAGTTCACACTGCTTTCAGAGGTTGGGGCTTCCAGTC 5244
301 AAGAGGTCCTCATTCACAGTACCAAGCCAGGTGCTGCAAGAGTCCCTGCGAGTCATGAAA 360
|||||
5245 AAGAGGTCCTCATTCACAGTACCAAGCCAGGTGCTGCAAGAGTCCCTGCGAGTCATGAAA 5304
361 CCAAGGAGGCTTGGGAAACCATCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
|||||
5305 CCAAGGAGGCTTGGGAAACCATCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 5364
421 GCGTGGGCTGGGCAAGGACCAAGTCTGAGTCAGAGCCAGAGGACGAGAGCTGATCCC 480
|||||
5365 GCGTGGGCTGGGCAAGGACCAAGTCTGAGTCAGAGCCAGAGGACGAGAGCTGATCCC 5424
481 AGCACTGCCCCGCGCTCTGCGATGCAATCTCTGCGCACTTGAGAACAGCTGTAGAG 540
|||||
5425 AGCACTGCCCCGCGCTCTGCGATGCAATCTCTGCGCACTTGAGAACAGCTGTAGAG 5484
541 AGGCACTGGCGCTCTTCCGACATTC 564
|||||
5485 AGGCACTGGCGCTCTTCCGACATTC 5508
|||||
```

RESULT 2
AC097722/c AC097722 150110 bp DNA linear HTG 24-JAN-2002

DEFINITION Homo sapiens chromosome UNK clone Rpl1-186D19, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

AC097722
AC097722

AC097722.2 GI:18308824
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 150110)
Waterson, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 150110)
Waterson, R.H.
Direct Submission
Submitted (22-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 24, 2002 this sequence version replaced gi:16306472.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: H.NH0186D19

----- Summary Statistics -----

Sequencing vector: M13; 54%
Sequencing vector: plasmid; 46%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 147449 bases at least Q40
Consensus quality: 148263 bases at least Q30
Consensus quality: 148688 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 150429; sum-of-contigs

Quality coverage: 8.73 in Q20 bases; agarose-fp
Quality coverage: 7.35 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 3944: contig of 3944 bp in length
* 3945 4044: gap of unknown length
* 4045 9798: contig of 5754 bp in length
* 9799 9898: gap of unknown length
* 9899 14019: contig of 4121 bp in length
* 14020 14119: gap of unknown length
* 14120 19392: contig of 5273 bp in length
* 19393 19492: gap of unknown length
* 19493 24872: contig of 5380 bp in length
* 24873 24972: gap of unknown length
* 24973 43958: contig of 18986 bp in length
* 43959 44058: gap of unknown length
* 44059 68992: contig of 24934 bp in length
* 68993 104669: gap of unknown length
* 104670 104769: contig of 35577 bp in length
* 104770 150110: contig of 45341 bp in length.
```

FEATURES

source

1..150110

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="UNK"

/clone="Rpl1-186D19"

1..3944

/note="assembly_name:Contig25"

3945..4044

/estimated_length=unknown

misc_feature

gap

```
misc_feature      4045. .9798  
/note="assembly_name:Contig26"  
gap              9799. .9898  
/estimated_length=unknown  
misc_feature      9899. .14019  
/note="assembly_name:Contig27"  
gap              14020. .14119  
/estimated_length=unknown  
misc_feature      14120. .19392  
/note="assembly_name:Contig28"  
gap              19393. .19492  
/estimated_length=unknown  
misc_feature      19493. .24872  
/note="assembly_name:Contig29"  
gap              24873. .24972  
/estimated_length=unknown  
misc_feature      24973. .43558  
/note="assembly_name:Contig30"  
gap              43958. .44058  
/estimated_length=unknown  
misc_feature      44059. .68392  
/note="assembly_name:Contig31"  
gap              68993. .69092  
/estimated_length=unknown  
misc_feature      69093. .104669  
/note="assembly_name:Contig32"  
clone_end:r7  
vector_side:left  
gap              104670. .104769  
/estimated_length=unknown  
misc_feature      104770. .150110  
/note="assembly_name:Contig33"  
clone_end:SP6  
vector_side:right
```

Query Match	Similarity	97.8%	Score 551.4	DB 14	Length 150110
Best Local	Similarity	99.6%	Pred. No. 2.7e-120		
Matches	563	Conservative	0	Mismatches	1
				Indels	1
				Gaps	1
Qy	1	GCAGCCCTGCGCTCCAGAAAAACGAAAGGACGACGACACATCAACGGGACTGACCCCCCATGTC	60		
Dp	50478	GCAGCCCTGCGCTCCAGAAAAACGAAAGGACGACGACACATCAACGGGACTGACCCCCCATGTC	50411		
Qy	61	GCTGAGGTGAGGAGGAGCTCTCTGAGGACGAGCCACGAGCCCGTCAGTGGTGTGACGCGCA	120		
Dp	50418	GCTGAGGTGAGGAGGAGCTCTCTGAGGACGAGCCACGAGCCCGTCAGTGGTGTGACGCGCA	50355		
Qy	121	GGGGCTCTTGCCATGCTGTGGGACACAGGGGCTGCTATCAGCTTACTCTGACATGACATGATCC	180		
Dp	50358	GGGGCTCTTGCCATGCTGTGGGACACAGGGGCTGCTATCAGCTTACTCTGACATGACATGATCC	50299		
Qy	181	CTGGTGTCCAGGCTCTGGAAGCTGTGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTGTC	240		
Dp	50298	CTGGTGTCCAGGCTCTGGAAGTGTGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTGTGTC	50233		
Qy	241	GCCGGCCATGCGCCCCCAACGTTGACACTACGTCGCTTGTCAAGGGTTGGGGCTTCCAGTC	300		
Dp	50238	GCCGGCCATGCGCCCCCAACGTTGACACTACGTCGCTTGTCAAGGGTTGGGGCTTCCAGTC	50175		
Qy	301	ACAGGGTCCCATCCACGTACCGTACCGCCAGCTGTGCTGCAGAAAGTCCCTCGCAGTCATGAAA	360		
Dp	50178	ACAGGGTCCCATCCACGTACCGTACCGCCAGCTGTGCTGCAGAAAGTCCCTCGCAGTCATGAAA	50115		
Qy	361	CCAAAGGAGGCTTGGGAAACCAATCTGAAAGGCGATGGCTTTGATTTAGTGAGAGGGGTGG	420		
Dp	50118	CCAAAGGAGGCTTGGGAAACCAATCTGAAAGGCGATGGCTTTGATTTAGTGAGAGGGGTGG	50055		
Qy	421	GGCTGGGGCTGGGCAAGGCGACACAGGTCTGAGTCAAGACACAGAGCGAGAAAGCTGGTCCCC	480		
Dp	50058	GGCTGGGGCTGGGCAAGGCGACACAGGTCTGAGTCAAGACACAGAGCGAGAAAGCTGGTCCCC	49999		
Qy	481	A-GCACTGCCCGCGCGCTCTGCGATGTGAGTCTTCTGGCCACTGTGAAGACAGCTGTGAA	539		

DB	49998	AGGACATGCCCCGCGCTTCGCAGTCAGATCTCTCGGACACTGTAGAAACACCTGTAGA	49939
OY	540	GAGGACATGAGTCTTTCCGACTTC	564
DB	49938	GAGGACATGCGCTCTTCGACTTC	49914
RESULT 3			
LOCUS	AP006345/c		
DEFINITION	AP006345	150214 bp	DNA
ACCESSION	Homo sapiens genomic DNA, chromosome 11 clone:RP11-186D19, complete sequence.		
VERSION	AP006345		
KEYWORDS	AP006345.4	GI:71891788	
SOURCE	HTG.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,		
JOURNAL	Totoki,Y. and Sakaki,Y.		
REFERENCE	Homo sapiens genomic DNA		
AUTHORS	Published Only in Database (2003)		
TITLE	2 (bases 1 to 150214)		
JOURNAL	Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,		
REFERENCE	Totoki,Y. and Sakaki,Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (30-APR-2003) Masahisa Hattori, The Institute of Physical		
JOURNAL	and Chemical Research (RIKEN), Genomic Sciences Center (GSC);		
REFERENCE	1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan		
AUTHORS	(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,		
TITLE	Tel:81-45-503-9111, Fax:81-45-503-9170)		
JOURNAL	On Aug 5, 2005 this sequence version replaced gi:48290861.		
REFERENCE	Location/Qualifiers		
FEATURES	1.150214		
SOURCE	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q"		
	/clone="RP11-186D19"		
ORIGIN			
Query Match	97.8%;	Score 551.4;	DB 8;
Best Local Similarity	99.6%;	Pred. No. 2.7e-120;	
Matches 563;	Conservative 0;	Mismatch 1;	Indels 1;
	Gaps 1;		
OY	1	GCAGCCCTGCTCAGAAAACAGAAAGGACGACACACATCAGCGTACACCCCATGTG	60
DB	85889	GCAGCCCTGCTCAGAAAACAGAAAGGACGACACACATCAGCGTACACCCCATGTG	85830
OY	61	GCTGAGAGTAGGAGGACCTCTTAGGACAGGAGCCAGGCGTACAGTGGGTGACGCA	120
DB	85829	GCTGAGAGCGAAGGAGGACCTCTTAGGACAGGAGCCAGGCGTACAGTGGGTGACGCA	85770
OY	121	GGGGCTTTCATGTGTGGGACAGGGGCTGCATACAGCTTACTCAGTCAATTCAGTCC	180
DB	85769	GGGGCTTTCATGTGTGGGACAGGGGCTGCATACAGCTTACTCAGTCAATTCAGTCC	85710
OY	181	CTGGTTCGACGCTCTTGAAGTCTGGAAGTGAGCAATGTTTCCATTAAAGAAAGTGTG	240
DB	85709	CTGGTTCGACGCTCTTGAAGTCTGGAAGTGAGCAATGTTTCCATTAAAGAAAGTGTG	85650
OY	241	GCGGACCATGCCCCCAAGCTTGACACATCATGCGCTTTGACAGGGTGGGGCTTCAAGTC	300
DB	85649	GCGGACCATGCCCCCAAGCTTGACACATCATGCGCTTTGACAGGGTGGGGCTTCAAGTC	85590
OY	301	ACAGGGTCCCATCAGCTACAGCCACAGCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAGAA	360
DB	85589	ACAGGGTCCCATCAGCTACAGCCACAGCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAGAA	85530

```
QY 361 CCAAGGAGAGGCTTGGAAAACAATCTGAAGGCGATGCTTTGATTAGTAGAGAGGCTGG 420
DB 8529 CCAAGGAGAGGCTTGGAAAACAATCTGAAGGCGATGCTTTGATTAGTAGAGAGGCTGG 85470
QY 421 GGCTGGGCTGGGCAAGGCCACAGGCTCTAGTCACAGCCAGAGGCAAGGCTGTCCCC 480
DB 85469 GGCTGGGCTGGGCAAGGCCACAGGCTCTAGTCACAGCCAGAGGCAAGGCTGTCCCC 85410
QY 481 A-GCACTGCCCCCGCTCTGCGATGCACTCTCTGCGCCACTGTAGAACACCTGTAGA 539
DB 85409 AGGCACTGCCCCCGCTCTGCGATGCACTCTCTGCGCCACTGTAGAACACCTGTAGA 85350
QY 540 GAGGCAGTGGCGCTCTTTCGGAAGCTTTC 564
DB 85349 GAGGCAGTGGCGCTCTTTCGGAAGCTTTC 85325

RESULT 4
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingberoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Porter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingberoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source 1..125020
organism="Homo sapiens"
mol_type="genomic DNA"
db_xref="taxon:9606"
chromosome="16"
map="16q24.3; between D16S520 and WI-12410"
note="Isolated from a patient with Huntington's
disease-like 2 (HDL2)"
complement(135581..35746)
repeat_region
rpt_type=tandem
rpt_unit="ctg"
complement(<36507..>36887)
gene/gene="JPH3"
note="synonym: Jp3"
complement(<36507..>36887)
product="junctophilin 3"
complement(<36507..>36887)
gene="JPH3"
note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
codon_start=1
product="junctophilin 3"
protein_id="AAL40941.1"
/db_xref="GI:17646245"
translation="MSSGGRFNFDDGGSGYCGMWDGKAKHGVCCTPGKGGGYTGSMS
HGFVLYGTWPSGNTYQGTMAQGRHGLGSKKRWYKGMWTHGFGRGVRECAQ
NGAKYEGTWSNGLQDGYGTETYSDG"
```

```
Query Match 7.6%; Score 43; DB 8; Length 125020;
Best Local Similarity 12.0%; Pred. No. 6.2;
Matches 64; Conservative 221; Mismatches 245; Indels 2; Gaps 1;

QY 23 AAGACGAGACACACTACCGGTGACTACCCCCCATGTGGCTGAGAGTAGAGGAGGCTCCT 82
DB 17541 MRSSKSSSMWSMBSMRSSWCMWAGYRRSKRSAGCAWBSKGRSTGMCRACKSKTGS 17482
QY 83 GAGGACAGGCGCAGGCGCAGCCCTCAGGTGGTGAAGCGAGGAGGTCTTGCATGTGGGAC 142
DB 17461 YGSTRSMKKKKKYSKYSRKGKKKCTCYCMKKYKTKKYSKMWYKMSWGYKRYKRC 17422
QY 143 AAGGCTGCATATACAGCTTACTCACTGACATCGATCGCTGCTGACCTCTGGAAGTC 202
DB 17421 MKKKGCTYRGMSSKSGSKSGKMRGSSYSTSKWSKWSKMKCMKRYMKRRRMR 17362
QY 203 TGAAGTGAACAATGTTCCATTAAAGAAAGTGTGGCCGACATCCCCCAAGCTT 262
DB 17361 GMSKSGMRGYAGRGYSSSSWMSSTRKRKSKYSYKRGKRGKMGWGMRRGSKYWSM 17302
QY 263 GCACACTGACTGCTTTCGAGGTTGGGCTTCCAGTCACAGGCTCCATCCACGTACCA 322
DB 17301 KMKRSMSSKCYSTKYSYSGRRKSKGWRSTKSKAKSSMRAGSKCTGSSYWSNNRNN 17242
QY 323 GCCCAGGTGGCTGAGAAAGT--CCCTGCAGTCATGAAACCAAGAGGCTTGGAAAC 380
DB 17241 RMGKTGCTMYRRBRAMNNGNAAGCTTCCCANNTNGGGGAAAGGCGSASRASCY 17182
QY 381 CACATCTGAAGGCAAGCTTGTATTAGTAGAGAGGTGGGCTGGCTGGGCAAGGCCA 440
DB 17181 KGRSSSKCYSGTTRRRKMKCKRRGSKGKMGWTRGSGGKTSYANGSRGCTYCMGW 17122
QY 441 CCAGTCTGAGTCAGAGCAGCAGGACGAGAACTGTGCCACACTGCCCCAGCACTGCCCGCTGTG 500
DB 17121 GRKGRCKMSKMKMYKYSYRKMKMTCKMKCTSYMAAMYRSMCMCKSCCGCTYMG 17062
QY 501 CGATGCACTGCTCTCTGGCAGCAGTGAACAACACTGTGATGAGAGGAGTGGCGGT 552
DB 17061 MSSYSGTYSKYSWMSKYSWMSRYSKSTSKAMRSSKRMGMGTGGRYGGGRSY 17010

RESULT 5
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingberoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Porter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingberoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source 1..125020
organism="Homo sapiens"
mol_type="genomic DNA"
db_xref="taxon:9606"
chromosome="16"
```


/map="16q24.3, between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-like 2 (HDL2)"
repeat_region complement(35581..35746)
/rpt_type=tandem
/rpt_unit="ctg"
gene complement(<36507..36887)
/gene="JPH3"
/note="synonym: JP3"
complement(<36507..36887)
/gene="JPH3"
/product="junctional protein 3"
complement(<36507..36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctional protein 3"
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translation="MSGGRFNPDDGSGYCGMEDKAKHGHVCTGPKGQGEYTSWS
HGFEVLYTTPSGNTYGTGTAAGCRRHIGLESKQKMYKGEWTHGFKRGVRECAQ
NGAKYEGTWSNGLDGYCTERYSDG"

ORIGIN

Query Match 7.3%; Score 41.4; DB 8; Length 125020;
Best Local Similarity 11.5%; Pred. No.15;
Matches 45; Conservative 169; Mismatches 176; Indels 0; Gaps 0;

OY 32 GCACACTACGCTGACTACACCCCATGTGGCTGTGAGGAGGAGCCTCTGAGCAGAGG 91
DB 51746 KSAYSQMTSMKXKCKSKYCAKSGCYKXKYSWYTSRSMWYTSYSCMYMSMMWSY 51805
OY 92 CCAGGAGCAGCCTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 151
DB 51806 MSWSKSGYTSCKMKCMCMSCMYRSARCTMRSTGTCYMSRSTGMSKMGKGCY 51865
OY 152 ATACAGCTTACTCACTGACATGAGTCCCTGTGCTGAGCAGCCTCTGAGAGTGTGAGA 211
DB 51866 RRSCKSMKTKGTCCKKSKSMKCATTMKSKSKSKMSYCGMCRASCMGCKYKMSMT 51925
OY 212 GCAATGTTTCCATTAGAAAGTGTGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 271
DB 51926 CMMSYTTTCTCYCYCYCMKSKYKRSYSYKSKSKMTYKSMRGSYRGRYSY 51985
OY 272 CTGCTTTTGGAGGTTGGGCTTCCAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 331
DB 51986 RYRMKCMCMGWTTRSKSKSKSYSSNSGMMWGKGTYSYMCASSYASCKMRMSATSWGCC 52045
OY 332 GCTGCAAGAGTCTCTGCAAGTCAATGAACCAAGGAGGAGGAGGAGGAGGAGGAGG 391
DB 52046 YMSYRSASCMGWSSTMSYSTRMSCHMSCHMSYSTRMSCHMSCHMSCHMSCHMSCH 52105
OY 392 GGCATGCTTGTATTAGTAGAGAGGAGGAGG 421
DB 52106 RGSWMGKSGMTKSKSGTGKGMWGRGSGWS 52135

RESULT 6
AC145527/c 320658 bp DNA linear HTG 19-JUL-2003
LOCUS AC145527/c

DEFINITION Atelestix albiventris clone LB4-283G23, WORKING DRAFT SEQUENCE, 21
unordered pieces.

ACCESSION AC145527
VERSION AC145527.1 GI:3296768
KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT.
SOURCE Atelestix albiventris (middle-African hedgehog)

ORGANISM Atelestix albiventris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Etriacidae;
Etriacinae; Atelestix.
REFERENCE 1 (bases 1 to 320658)
AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,

TITLE Peng,Z., Malinov,I. and Rubin,E.M.
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 320658)
AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
COMMENT Draft Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: PGABERK
Center Project Name: E024-283G23
Bac Clone Name: LB4-283G23

Additional information on comparative analysis and ordering are
available at:
http://pga.lbl.gov/cgi-bin/search_cvs.cgi?type=nt&value=SRBFP1
Funding agent: Programs for Genomic Applications (NHLBI)
if library name is LBI to LB4, please see website
for a description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraries.htm> as LBML-1 to LBML-4.
Summary Statistics:
Sequencing vector: Plasmid; pUC18
Chemistry: Dye-terminator Big Dye
Assembly Program: Phrap version 0.990329.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2325: contig of 2325 bp in length
2326 2425: gap of unknown length
2426 4760: contig of 2335 bp in length
4761 4860: gap of unknown length
4861 7844: contig of 2984 bp in length
7845 7944: gap of unknown length
7944 10691: contig of 2747 bp in length
10692 10791: gap of unknown length
10792 15758: contig of 4967 bp in length
15759 15858: gap of unknown length
15859 18758: contig of 2900 bp in length
18759 18858: gap of unknown length
18859 23515: contig of 4657 bp in length
23516 23615: gap of unknown length
23615 29133: contig of 5518 bp in length
29134 29233: gap of unknown length
29234 33870: contig of 4637 bp in length
33871 33970: gap of unknown length
33971 33971: contig of 5546 bp in length
33971 39516: gap of unknown length
39517 45912: contig of 6296 bp in length
45913 46012: gap of unknown length
46013 52370: contig of 6358 bp in length
52371 52470: gap of unknown length
52471 64704: contig of 12234 bp in length
64705 64804: gap of unknown length
64805 76264: contig of 11460 bp in length
76265 76365: gap of unknown length
76365 94758: contig of 18394 bp in length
94759 94858: gap of unknown length
94859 118713: contig of 23855 bp in length
118714 118813: gap of unknown length
118814 148021: contig of 29208 bp in length
148022 148121: gap of unknown length
148122 178939: contig of 30818 bp in length
178940 211754: gap of unknown length
211755 211755: contig of 32715 bp in length
211755 211854: gap of unknown length

FEATURES

source

211855 263200: contig of 51346 bp in length

263201 263300: gap of unknown length

263301 320658: contig of 57358 bp in length.

Location/Qualifiers

1..320658

/organism="Atelalex albiiventris"

/mol_type="genomic DNA"

/db_xref="taxon:9368"

/clone="LB4-283G23"

3226..2425

/estimated_length=unknown

4761..4860

/estimated_length=unknown

7845..7944

/estimated_length=unknown

10692..10791

/estimated_length=unknown

15759..15858

/estimated_length=unknown

18759..18858

/estimated_length=unknown

23516..23615

/estimated_length=unknown

29134..29233

/estimated_length=unknown

33871..33970

/estimated_length=unknown

39517..39616

/estimated_length=unknown

45913..46012

/estimated_length=unknown

52371..52470

/estimated_length=unknown

64705..64804

/estimated_length=unknown

76265..76364

/estimated_length=unknown

94759..94858

/estimated_length=unknown

118714..118813

/estimated_length=unknown

148022..148121

/estimated_length=unknown

178940..179039

/estimated_length=unknown

211755..211854

/estimated_length=unknown

263201..263300

/estimated_length=unknown

263201..263300

/estimated_length=unknown

ORIGIN

Query Match 7.1%; Score 40.2; DB 14; Length 320658;

Best Local Similarity 58.0%; Pred. No. 25;

Matches 91; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 1 GCAGCCCTCTCAGAAAACAGAGGACGACACACTCAGCGTACTGACCCCATGNG 60

DB 316431 GACCCCACTTGCACAAACAGAGCGTGGGCCAC---ACACACTCAGCGTCTTGG 316375

QY 61 GCTGAGGTGAGGAGGACCTCTCTGAGGACGAGGCGGCGTCAAGTGGGTGACGGCA 120

DB 316374 TGCCTTGGGCTGTGGCGCCGCGGAGCTGTGCTCCAGCAACGGGTGAGTGGGGC 316315

QY 121 GGGGCTTGGCATGTGTGGGACAGAGGGCTGCATACG 157

DB 316314 TGGCTTCCGCGCATCAGGCGCACAGTGGCTTATGAG 316278

RESULT 7

AC163209/c AC163209 153129 bp DNA linear ROD 03-AUG-2005

LOCUS

DEFINITION Mus musculus chromosome 3, clone RP24-11708, complete sequence.

ACCESSION AC163209

VERSION

AC163209.2 GI:71725541

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Birtten,B., Nusbaum,C. and Lander,E.

TITLE

Mus musculus chromosome 3, clone RP24-11708

JOURNAL

Unpublished

REFERENCE

AUTHORS

2 (bases 1 to 153129)

Birtten,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Cornu,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., Fitzerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (07-JUN-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 153129)

Birtten,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Cornu,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., Fitzerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (03-AUG-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Aug 3, 2005 this sequence version replaced gi:67003668.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submission@broad.mit.edu

----- Project Information

Center project name: U32928

Center clone name: 117_O_8

```
FEATURES
Source
-----
Location/Qualifiers
1..153129
/orrganism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP24-11708"
/clone_11b="RPCI-24 Male Mouse BAC"
repeat_region
2843..2893
/rpc_family="AT-rich"
complement(2985..3180)
repeat_region
5441..5470
/rpc_family="B4"
repeat_region
5748..5876
/rpc_family="(GAA)n"
repeat_region
6984..7110
/rpc_family="(TCTA)n"
repeat_region
7462..7801
/rpc_family="MLT1C"
complement(9528..9910)
repeat_region
9937..10085
/rpc_family="MT2B"
repeat_region
10825..10982
/rpc_family="MERSB"
repeat_region
complement(11051..11261)
/rpc_family="(TG)n"
repeat_region
12012..12033
/rpc_family="B4"
repeat_region
12907..13123
/rpc_family="(A)n"
repeat_region
13453..13812
/rpc_family="MLT1C"
repeat_region
complement(14278..14503)
/rpc_family="MTE"
repeat_region
complement(14517..14746)
/rpc_family="L1"
repeat_region
14745..15433
/rpc_family="L1_MM"
repeat_region
complement(15710..15820)
/rpc_family="L1MA4A"
repeat_region
complement(15823..16050)
/rpc_family="L1MA4A"
repeat_region
16057..16122
/rpc_family="(TG)n"
complement(16123..16262)
/rpc_family="B3A"
repeat_region
16576..16671
/rpc_family="(TG)n"
repeat_region
16676..16856
/rpc_family="(TATATG)n"
repeat_region
16862..16925
/rpc_family="(TATG)n"
repeat_region
16928..17091
/rpc_family="(TATATG)n"
complement(17311..17610)
/rpc_family="MLT1A0"
complement(17617..19453)
/rpc_family="L1_MM"
18133..18163
/notes="single clone coverage"
repeat_region
19454..20704
/rpc_family="L1"
repeat_region
20758..20824
/rpc_family="MIR"
complement(20852..21030)
/rpc_family="MIR"
repeat_region
21915..21942
/rpc_family="(TTTGTG)n"
complement(21943..22157)
repeat_region
```

```
repeat_region
22301..22397
/rpc_family="B3"
repeat_region
complement(22732..22787)
/rpc_family="L2"
repeat_region
22786..23404
/rpc_family="L1"
repeat_region
complement(23405..23879)
/rpc_family="L1_MM"
repeat_region
23880..24023
/rpc_family="MERVL_LTR"
repeat_region
complement(24023..24051)
/rpc_family="L1_MM"
repeat_region
24051..24205
/rpc_family="(TCTCC)n"
repeat_region
25614..25638
/rpc_family="AT-rich"
repeat_region
26685..26707
/rpc_family="(TTTG)n"
repeat_region
28871..28954
/rpc_family="(TA)n"
repeat_region
complement(29240..29627)
/rpc_family="TAPRY_LTR"
repeat_region
complement(29660..29993)
/rpc_family="ORR1D"
repeat_region
30197..30223
/rpc_family="(TTC)n"
repeat_region
complement(30226..30374)
/rpc_family="B1_MM"
repeat_region
30733..30785
/rpc_family="(TTC)n"
repeat_region
complement(30891..31095)
/rpc_family="MIR"
repeat_region
complement(30926..31037)
/notes="single clone coverage"
31116..31151
/rpc_family="(GGGGA)n"
repeat_region
31130..31168
/notes="single clone coverage"
31345..31410
/rpc_family="(TG)n"
repeat_region
complement(31417..31571)
/rpc_family="RSINE1"
repeat_region
complement(33166..33865)
/rpc_family="RMR6A"
38071..38261
/rpc_family="B3A"
repeat_region
38280..38428
/rpc_family="B3A"
repeat_region
38402..38558
/rpc_family="(CA)n"
39152..39203
/rpc_family="(TTR)n"
repeat_region
```

Query Match 7.1%; Score 40; DB 9; Length 153129;
Best Local Similarity 52.4%; Pred. No. 31;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
QY 301 ACAGGTCATCACTACCAAGCCAGGTGCTGCAGAAAGTCTCTCGATCATGAA 360
DB 76180 ACAAGTAGCTGCTGGAGACGAGGCGAGGCGAGGCGCACTGACGAGCCAG 76121
QY 361 CCAAGGAGGCTTGGGAAACCATCTGAAGGGCATGGCTTTGATTAGTAGAGGGTG 420
DB 76120 GACAGGAGCCAGGGGTGAGACTTAGGTGAGGCGAGGCGCCAGAGGCGAGG 76061
QY 421 GGCTGGCTGGCAAGGCCACCAAGTCTGATCAGAGCCAGAGCGAGG 468
DB 76060 GGGGGGCGCAGGCGCAGAGGCGAGGCGAGGCGCAGAGGCGAGG 76013
```

RESULT 8
AC111047 230041 bp DNA linear ROD 10-OCT-2003
LOCUS AC111047
DEFINITION Mus musculus chromosome 3, clone RP23-176H24, complete sequence.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	TITLE AUTHORS JOURNAL REFERENCE
AC111047 AC111047.9 GI:37620322 HTG. Mus musculus (house mouse) Mus musculus Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Euarctontoglires; Glires; Rodentia; Sciuromorphi; Muridae; Muridae; Murinae; Mus.	REFERENCE AUTHORS TITLE JOURNAL REFERENCE 2 (bases 1 to 230041) Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bokhgalter, B., Brown, A., Camarata, J., Campolongo, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gold, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Kartas, A., Kelle, C., Larocque, K., Lamazares, R., Lander, T., Lechoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roesti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Struss, N., Sudirmanlan, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
REFERENCE AUTHORS JOURNAL REFERENCE	REFERENCE AUTHORS TITLE JOURNAL REFERENCE 3 (bases 1 to 230041) Birren, B., Nusbaum, C. and Lander, E. Submitted (18-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE AUTHORS JOURNAL REFERENCE	REFERENCE AUTHORS TITLE JOURNAL REFERENCE 4 (bases 1 to 230041) Birren, B., Muehlbauer, C. and Lander, E. Submitted (07-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE AUTHORS JOURNAL REFERENCE	REFERENCE AUTHORS TITLE JOURNAL REFERENCE Submitted (10-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Oct 10, 2003 this sequence version replaced gi:34495134. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center

Center code: MIBR	
Web site: http://www-seq.wi.mit.edu	
Contact: sequence_submissions@genome.wi.mit.edu	
----- Project Information	
Center project name: L19822	
Center clone name: 176_H_24	

FEATURES	
source	Location/Qualifiers
	1. .231041
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="3"
	/map="3"
	/clone="RP23-176H24"
	/clone_11b="RPC1-23 Female Mouse BAC"
misc_feature	1. .21291
	/note="wgs and extension
	clone_end:5p6"
repeat_region	1144. .1290
	/rpt_family="B1_MM"
repeat_region	2418. .2629
	/rpt_family="L2"
repeat_region	2641. .2696
	/rpt_family="MER91A"
repeat_region	3023. .3043
	/rpt_family="TCTG)n"
repeat_region	3086. .3126
	/rpt_family="(CA)n"
repeat_region	3126. .3176
	/rpt_family="(GA)n"
repeat_region	3299. .3449
	/rpt_family="L2"
unsure	3572. .3576
	/note="<30 qual SINGL region"
unsure	complement(5063..5383)
	/note="single clone coverage"
unsure	5464. .5470
	/note="<30 qual SINGL region"
unsure	5498. .5503
	/note="<30 qual SINGL region"
unsure	5539. .5544
	/note="<30 qual SINGL region"
unsure	5564. .5568
	/note="<30 qual SINGL region"
repeat_region	complement(5599..6129)
	/rpt_family="Lx2B"
repeat_region	6272. .6332
	/rpt_family="(GA)n"
repeat_region	7692. .7878
	/rpt_family="ORR1B1"
repeat_region	complement(8072..8141)
	/rpt_family="MIR"
repeat_region	8906. .9058
	/rpt_family="BA4"
repeat_region	9622. .9758
	/rpt_family="(TCC)n"
repeat_region	complement(10521..10594)
	/rpt_family="Zaphod"
repeat_region	complement(12702..12858)
	/rpt_family="B3"
repeat_region	complement(12871..13230)
	/rpt_family="MTD"
repeat_region	13255. .13275
	/rpt_family="AT_rich"
repeat_region	13286. .14525
	/rpt_family="single clone coverage"
unsure	14031. .14049
	/rpt_family="Lx4"
repeat_region	14535. .14580
	/rpt_family="(CAAT)n"
repeat_region	complement(14607..14691)
	/rpt_family="Lx9"

```

repeat_region 14812..14846
/rpc_family="CMAAA)n"
repeat_region 15203..15348
/rpc_family="B3A"
repeat_region 15345..15391
/rpc_family="(CA)n"
repeat_region 15392..15437
/rpc_family="B3A"
repeat_region 15846..15975
/rpc_family="MTD"
repeat_region 16110..16376
/rpc_family="MTC"
repeat_region 18724..18891
/rpc_family="L1MB8"
repeat_region 19356..19561
/rpc_family="B3"
repeat_region 19961..19988
/rpc_family="AT-rich"
repeat_region 19990..20035
/rpc_family="(TCTA)n"
repeat_region 20124..20180
/rpc_family="AT-rich"
repeat_region 20343..20479
/rpc_family="B1-MW"
repeat_region 20694..21089
/rpc_family="MTA"
misc_feature 21292..21297
/Note="clone_boundary
clone_end:SP6
site:ECORI"
repeat_region 21395..21487
/rpc_family="(CAGGG)n"
repeat_region 21517..21672
/rpc_family="L1MB8"
repeat_region 27914..28115
/rpc_family="(UR1A"
repeat_region 30025..30065
/rpc_family="(TG)n"
repeat_region 30216..30311
/rpc_family="RSINR1"
repeat_region 30611
/rpc_family="B4"
repeat_region 31033..31131
/rpc_family="(CCA)n"
repeat_region 32917..33316
/rpc_family="(TG)n"
repeat_region 33334..33515
/rpc_family="(TG)n"
repeat_region 33712..33732
/rpc_family="AT-rich"

Query Match 7.1%; Score 40; DB 9; Length 230041;
Best Local Similarity 52.4%; Pred. No. 29;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 301 ACAGGCTCCATCCACGTAACAGCCAGCTGCGACAGAGTCTCCGCACTATGAA 360
Db 21304 ACAAGTGTCTGTGGAGCCAGGCGACAGGCGCGGCGCCAGAGCAAGAGCCAA 21363
QY 361 CCAAGGAGGCTGTGGGAACACATCTGAAGGCACTGCTTTGATTTTGTAGAGGCTG 420
Db 21364 GACAGGAGCCAGGCGGTGAGCTAGGGTGTGAGGCGAGGCGCGGCGCCAGAGCAAGAGCCAGG 21423
QY 421 GGCTGTGGGCAAGGCCACAGGCTGTAGTCAAGAGCCAGAGGCAAG 468
Db 21424 GGCAGGCGCAGGCGAGGCGAGGCGCGAGAGCCAGGAGAGGCAAG 21471

RESULT 9
AC067978
LOCUS AC067978 103034 bp DNA linear HTG 03-JUL-2000

```

```

DEFINITION Homo sapiens chromosome 2 clone CTD-2350K7 map 2, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC067978
VERSION AC067978.2 GI:8901271
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 103034)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 2, clone CTD-2350K7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 103034)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E.,
Anderson,S., Baldwin,D., Barna,N., Baetsen,V., Bada,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choquet,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Doming,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,D., Gardyna,S., Ginde,S., Goylette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,D.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Labocque,K., Lamazares,R., Landers,T., Lehoczy,D.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meidirim,D., Meneses,L., Minova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefley,S., Theodore,D., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Videl,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 3, 2000 this sequence version replaced gi:7656720.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9518
Center clone name: 2350_K7
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 9545 bases at least Q40
Consensus quality: 9946 bases at least Q30
Consensus quality: 100858 bases at least Q20
Insert size: 98000; agarose-fp
Quality coverage: 4.5 in Q20 bases; sum-of-coverage
Quality coverage: 4.4 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1585: contig of 1585 bp in length
* 1586 1685: gap of 100 bp

```


LOCUS AC007362 116039 bp DNA linear PRI 16-APR-2005
DEFINITION Homo sapiens BAC clone RP11-150F11 from 2, complete sequence.
ACCESSION AC007362
VERSION AC007362.4 GI:10801453
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 116039)
AUTHORS Bemis,G., Kallick,J., Maupin,R. and Davidson,T.
TITLE The sequence of Homo sapiens BAC clone RP11-150F11
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 116039)
AUTHORS Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (23-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 116039)
AUTHORS Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (15-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 116039)
AUTHORS Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (25-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 116039)
AUTHORS Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (29-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 116039)
AUTHORS Waterston,R.
TITLE Direct Submision
JOURNAL Submitted (14-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 7 (bases 1 to 116039)
AUTHORS Wilson,R.K.
TITLE Direct Submision
JOURNAL Submitted (16-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 14, 2000 this sequence version replaced gi:9211352.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watsn.wustl.edu
----- Summary Statistics
Center project name: H_NH0150F11

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E.,
Tareno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-325M10, 200 base pair
overlap; the clone sequenced to the right is RP11-394E1, 200 base
pair overlap. Actual start of this clone is at base position
117994 of RP11-325M10; actual end is at base position 115845 of
RP11-150F11.

FEATURES
source
Location/Qualifiers
1..116039
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-150F11"
/clone_11b="RPc1-11"
12630..13592
/note="CpG island (%GC=70.3, o/e=0.84, #CpGs=108)"
14041..97553
/gene="NRP2"
join(14041..14113,28294..28471,46993..47124,53228..53458,
54535..54690,56663..56832,58641..58796,71269..71413,
73953..74302,76496..76640,80475..80591,83585..83725,
94424..94686,96224..96320,97553..97553)
/gene="NRP2"
join(14041..14113,28294..28471,46993..47124,53228..53458,
54535..54690,56663..56832,58641..58796,71269..71413,
73953..74302,76496..76640,80475..80591,83585..83725,
94424..94686,96224..96320,97553..97553)
/gene="NRP2"
/note="Homo sapiens neuropilin 2 (NRP2), mRNA.;
H_NH0150F11.1
This gene was based on gi(4505458)
Continues as H_NH0394E01.1"
/codon_start=1
/product="unknown"
/protein_id="AA03216.1"
/db_xref="GI:62702291"
/translation="MDMPPLTWFLALVFSHQVQGQDPDPCGRLSKAGYITSPG
YPDYPSHONCEWVVAPEPNQKIVLNFNPFELKDKCKYFELIRGDESDLLG
KHGKNIAPTIISGSMLYIRFTSDVARQGFSLRYEIFYTGSDCKNFTSPNGTI
ESGPPEKYPNLLCTFTILAKPMKEIILQFLFDLEHPLDVGEGDCKYMDLWDG
IPIVGLIKGYCGKTPSELRSNGIISLTFHTMAVAKDGFSAVYVHDEPLEPO
CNVPLGMSGRRIANEQISASTSYSDGKRTPOQSILHGDNDKWTNRLDSNKXYLOYDRL
FLTULATATGALISRETONGYVYKSYLSTNGEDMVRHGNLGNLPOANDATL
VLNLKHALPLTRFVRIRPQTMHSGIALRLFCGRVYDACSNNLSQLINDSQI
SASSTQELTWSAARLVSRSGWFPRIPOAPGEBMWLOVLTGPTKYGVYIIGARG
GSTITAVEARAFVRKFRVSYSLNKDMEYIDDPQOQKLFEGNMHYTPPIRRDPI
PAQYVRYVPERWSPAGTGMRLGVLCGMDTSDCKPVTETIGPTVKSSEETTPVTESEAT
EGGNCSTFEDDDXLDLPSSGFCNDFLBBPCCGMYDAKMLRTTMASSSPDRDFPD
DRNFLRLQSDSQRGGYARLISPVHLPRSPVCMFQYQATGKGVLAQVREASQES
KLVLVIREDDCGEGWKHGRILLPSYDMEYQIVFEGVIGKRGSGIALIDDIRISTDVEE
NCMPISAFIA"
15626..17342
/gene="NRP2"
/note="CpG island (%GC=64.9, o/e=0.68, #CpGs=145)"

ORIGIN
Query Match 7.1%; Score 39.8; DB 8; Length 116039;
Best Local Similarity 64.8%; Pred. No.36;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
4 GCCCTGCTCAGAAACGAGAGCGACGACACTCAGTGACTCACCCCATGTGCT 63
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 77259 GCCCTCCCTCGGAACGAGAGGCTGGACCTCCGTGGGCTCACCCTTCATTGGAA 77200

64 GGAGGTGAGGAGCCTCTGTGAGCGAGGCCA 94

Db 77199 TGATGAGAGGTGTCTCCCAAGGCTGGGCCA 77169

RESULT 12
AC159874/c
LOCUS
DEFINITION Bos taurus clone CH240-5418, *** SEQUENCING IN PROGRESS ***, 26
unordered pieces.

AC159874 233561 bp DNA linear HTG 01-JUL-2005
AC159874.3 GI:68267115
HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
Bos taurus (cow)
Bos taurus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS
Munry, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, K., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, J.,
Devila, M. L., Davis, S., Davy-Carroll, L., De Anda, C., Dederich, D.,
Dejgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Freese, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunnarsson, P., Haland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, J., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovacs, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangun, A.,
Mangun, B., Mapua, P., Martin, K., Martin, R., Martine, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muidada, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nweokwelen, O., Okwona, G., Olampunsgoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plazzer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C. D., Smaja, D.,
Snead, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Syatel, A., Tabors, Z., Umani, K.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, S., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G., and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 233561)
Direct Submission
Unpublished

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Worley, K. C.
Direct Submission
Submitted (13-APR-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 233561)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:62543289.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FDDS
Center clone name: CH240-5418
----- Summary Statistics
Assembly program: Atlas 3.0.
Consensus quality: 226744 bases at least Q40
Consensus quality: 228422 bases at least Q30
Consensus quality: 229869 bases at least Q20
Estimated insert size: 229962; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 26 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1
4476 4475: contig of 4475 bp in length
4476 4794: gap of 319 bp
4795 10004: contig of 5210 bp in length
10005 10054: gap of 50 bp
10055 14989: contig of 4935 bp in length
14990 15039: gap of 50 bp
15040 23192: contig of 8153 bp in length
23193 23242: gap of 50 bp
23243 27656: contig of 4414 bp in length
27657 27706: gap of 50 bp
27707 32345: contig of 4639 bp in length
32346 32395: gap of 50 bp
32396 80065: contig of 47670 bp in length
80066 80258: gap of 193 bp
80259 95526: contig of 15268 bp in length
95527 95576: gap of 50 bp
95577 11151: contig of 15575 bp in length
11152 11201: gap of 50 bp
11202 127875: contig of 16674 bp in length
127876 127925: gap of 50 bp
127926 130975: contig of 3050 bp in length
130976 136494: gap of 89 bp
136494 136494: contig of 5430 bp in length

* 136495 136594: gap of unknown length
* 136595 140341: contig of 3747 bp in length
* 140342 140391: gap of 50 bp
* 140392 141698: contig of 1307 bp in length
* 141699 141748: gap of 50 bp
* 141749 151343: contig of 9595 bp in length
* 151344 151393: gap of 50 bp
* 151394 154790: contig of 3397 bp in length
* 154791 154840: gap of 50 bp
* 154841 166899: contig of 12059 bp in length
* 166900 172623: contig of 5674 bp in length
* 172624 172673: gap of 50 bp
* 172674 185758: contig of 13685 bp in length
* 185759 185808: gap of 50 bp
* 185809 210143: contig of 2435 bp in length
* 210144 210193: gap of 50 bp
* 210194 226003: contig of 15810 bp in length
* 226004 226103: gap of unknown length
* 226104 227673: contig of 1570 bp in length
* 227674 227773: gap of unknown length
* 227774 229134: contig of 1361 bp in length
* 229135 229234: gap of unknown length
* 229235 230327: contig of 1093 bp in length
* 230328 230427: gap of unknown length
* 230428 231641: contig of 1214 bp in length
* 231642 231741: gap of unknown length
* 231742 233561: contig of 1820 bp in length.

FEATURES
source
1..233561
Location/Qualifiers

gap /organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-5418"
4476..4794
/estimated_length=319
gap 10005..10054
/estimated_length=50
gap 14990..15039
/estimated_length=50
gap 23193..23242
/estimated_length=50
gap 27657..27706
/estimated_length=50
gap 32346..32395
/estimated_length=50
gap 80066..80258
/estimated_length=193
gap 95527..95576
/estimated_length=50
gap 11152..11201
/estimated_length=50
gap 127876..127925
/estimated_length=50
gap 130976..131064
/estimated_length=89
gap 136495..136594
/estimated_length=unknown
gap 140342..140391
/estimated_length=50
gap 141699..141748
/estimated_length=50
gap 151344..151393

Query Match 7.1%; Score 39.8; DB 14; Length 233561;
Best Local Similarity 60.7%; Pred. No. 33;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 274 GCCTTGACGGGTGGGCTTCACGACAGGGTCCACGATCCAGACGCCAGATGGC 333
DB 125658 GCTTGCCCGGGGATTAAGCTGCTGGCTGTGAGGCTGGCCCCCAGCTCCAGCACTGCCGC 125599
QY 334 TGACAGAGGTCCCTCGACATCATGAACCAAGGAGGCTTGGAAC 380

Db 125598 TCTGCAAGTCCCTCAGGTCATCTGTCAGAGTGGGGGTGGGAC 125552

RESULT 13
AC138582/C
LOCUS AC138582 119613 bp DNA linear PRI 04-AUG-2005
DEFINITION Pan troglodytes clone rp43-21p1, complete sequence.
AC138582
VERSION AC138582.22 GI:71067221
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.

REFERENCE 1 (bases 1 to 119613)
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Pan troglodytes BAC Clone rp43-21p1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 119613)
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 3 (bases 1 to 119613)
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT ----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

On Jul 23, 2005 this sequence version replaced gi:67514672.

FEATURES
source
1..119613
Location/Qualifiers

gap /organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="rp43-21p1"
/clone_lib="RPC1 - 43 Male Chimpanzee BAC Library"

ORIGIN

Query Match 7.0%; Score 39.6; DB 8; Length 119613;
Best Local Similarity 46.9%; Pred. No. 40;
Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 3 AGCCCTGCTCAGAAACAGAAAGACGACGACACTCAGCGTGACTACCCCATGTGGC 62
DB 85272 AGCTGAGCAGCAACCTGGGAGGACCTGCTGGCTGTGATGGGATCAGAGG 85213
QY 63 TGAAGTGAAGAGCCTCTGAGGACGAGGACGAGGACCCCTCAGTGGGTGACGAGG 122
DB 85212 CCTAGTGAAGGAGGAGATGAGAGCCAGGCTTCAGAGGAGCAGGAGTGGGAG 85153
QY 123 GGTCTTGCATGTGGGACAGGGGCTGACATACAGCTTACTCAGTGAATCAGTCCCT 182
DB 85152 GTCATGGGCAAGTGGGAGGTCAAGCTGTGAGCACTTCCCAAGGCCAATGACCAAAAT 85093
QY 183 GGTGCAAGCTCTGGAATCTGGAAGTGAAGCAATGTTCCATTAAAGGAAGTGTGGC 242
DB 85092 GAGACAACTTTAAGACTGACAGAGCATGTGGGGGCTGTGGAGCCAGGAAGCTAAGTGC 85033

```

Oy      243 CGGCGATGGCCCCCAACGTTCG 264
Db      85032 ACACAACACCCCGAGCTCTGC 85011

RESULT 14
LOCUS   AC000029/c
DEFINITION Homo sapiens Chromosome 22q11.2 BAC Clone 865e9 In GNAZ-BCR Region, complete sequence.
ACCESSION AC000029
VERSION   AC000029.17
KEYWORDS  GTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS  Kim,U.-J. and Simon,M.
JOURNAL   Unpublished
REFERENCE
AUTHORS  Wang,Z. and Roe,B.A.
JOURNAL   Homo sapiens Chromosome 22q11.2 BAC Clone 865e9 In GNAZ-BCR Region
REFERENCE
AUTHORS  Roe,B.A.
JOURNAL   Direct Submission
TITLE     Submitted (21-OCT-1996) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE
AUTHORS  Roe,B.A.
JOURNAL   Direct Submission
TITLE     Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE
AUTHORS  Roe,B.A.
JOURNAL   Direct Submission
TITLE     Submitted (24-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE
AUTHORS  Roe,B.A.
JOURNAL   Direct Submission
TITLE     Submitted (01-JUN-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE
AUTHORS  Roe,B.A.
JOURNAL   Direct Submission
TITLE     Submitted (09-APR-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT   On Sep 24, 1999 this sequence version replaced gi:5870976. Because these overlapping clones came from different libraries.
FEATURES             location/Qualifiers
Source               1..142979
                     /organism="Homo sapiens"
                     /mol_type="Genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22q11.2"
                     /clone="865e9"

ORIGIN
Query Match       7.0%; Score 39.6; DB 8; Length 142979;
Beet Local Similarity 46.9%; Pred. NO.39;
Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
3 AGCCTGCTCAGAAGAGCGAAGCGACGACTCACGGTACTCACCCCATTGTCG 62

```

Db	133180	AGCTCAGACGAACTGGGGGAGGCCACCCTGTGCTGGCCGTGTGAATGGGATCAACAGAGA	133121
Qy	63	TGAGAGTGAAGGAGCGCTTCCTTAGGCAAGGGCCAGGGCAGCCGTTCAGTGCGGTGAACGGCAG	122
Db	133120	CCTAAGTACAGGAGGAGACTGTGAGCCAGGGCCCTCAAGAGAGATGGGGGTGGCATTTGGGGAG	133061
Qy	123	GGTCTTGCATATGTGTGGGACAGGGGCTGTGATPACAGTTACTAGTAACAATGAGTCCCT	182
Db	133060	GTCAATGGGCAAGTGTGGCAGGTCAAGGCTGTGAGCACCTTCCCAAGAGCCAATGACCAAAT	133001
Qy	183	GGTGCAGACCTCTGTGAAGTCTTGAAGTGAAGCATGTGTTCCATTAAAGAAAAGTGTGGC	242
Db	133000	GAGAACACTTTTGGGACTGACAGAGCATGTGGGGGCTCTGGAGCCAGAAAGCTAATGTGC	132941
Qy	243	CGGCATGCCCCCAACGTTGC	264
Db	132940	ACACAAACCCCAGGCTGTGC	132919
RESULT 15			
LOCUS	AC148588		
DEFINITION	AC148588	190018 bp	DNA linear HTG 23-JUL-2005
VERSION	AC148588.26	GI:71067215	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Pan.		
REFERENCE	1	(bases 1 to 190018)	
AUTHORS	Lau,C., Fujiyama,A. and Roe,B.A.		
TITLE	Pan troglodytes BAC Clone pcb-92F23		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 190018)	
AUTHORS	Lau,C., Fujiyama,A. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-MAR-2004) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
COMMENT	On Jul 23, 2005 this sequence version replaced gi:66473101.		
	----- Genome Center		
	Center: Department Of Chemistry And Biochemistry		
	The University Of Oklahoma		
	Center code:UOKNOR		

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 3 contigs. Gaps between the contigs		
	* are represented as runs of N. The order of the pieces		
	* is believed to be correct as given, however the sizes		
	* of the gaps between them are based on estimates that have		
	* provided by the submittor.		
	* This sequence will be replaced		
	* by the finished sequence as soon as it is available and		
	* the accession number will be preserved.		
*	1	137286:	contig of 137286 bp in length
*	137287	137386:	gap of unknown length
*	137387	145329:	contig of 7943 bp in length
*	145330	145429:	gap of unknown length
*	145430	190018:	contig of 44589 bp in length.
FEATURES			
SOURCE			
	1..190018		
	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		

/clone="ptb-94f23"
/clone_lib="PTB Male Chimpanzee BAC Library"
137287..137386
/estimated_length=unknown
145330..145428
/estimated_length=unknown

ORIGIN

Query Match 7.0%; Score 39.6; DB 14; Length 190018;
Best Local Similarity 46.9%; Pred. No. 38;
Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
QY      3 AGCCCTGCCTCAGAAAACGAAAGACGACACACTCAGGTGACTCACCCCATGTGGC 62
DB      165064 AGCTCAGCAGCAACCTGCGGAGGACCACCTGCTGGCTGTGAGTGGGATCACAGAGG 165123
QY      63 TGGAGGTAGGGAGCCTCTCTGAGGCAAGGCCAGGGCAGCCCTCAGGTGGGTACCGCAGG 122
DB      165124 CCTAGGTACGGAGGACTGAGGCAAGGCTCTGAGAGGAGCGGGGTGGCATTTGGGGAG 165183
QY      123 GGTCTTGCCATGTGGGACAGGGGCTGCAATACAGCTTACTCAGTGACAAATCGAGTCCCT 182
DB      165184 GTCATGGGCAAGGTGGGCAAGGTACAGGCTGTGAGCACTTCCCGAGGCCAATGACCAAAAT 165243
QY      183 GGTGCCAGCCTCTGAAAGTCTGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTGGC 242
DB      165244 GAGACAACCTTTAGACTGACAGACATGTGGGGGCTCTGGAGCCAGAAAGCTTAAGTGTG 165303
QY      243 CGGCCATGCCCCCAACGTTGC 264
DB      165304 ACACAACACCCCAAGCTCTGC 165325
```

Search completed: March 25, 2006, 17:51:44
Job time : 3011.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 15:57:57 ; Search time 472 Seconds
(without alignments)
7963.743 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagcctgcctcagaaac.....agtcgctcttcgacttc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_GeneSeq_21.*
1: GeneSeq1980s:*
2: GeneSeq1990s:*
3: GeneSeq2000s:*
4: GeneSeq2001s:*
5: GeneSeq2001bs:*
6: GeneSeq2002as:*
7: GeneSeq2002bs:*
8: GeneSeq2003as:*
9: GeneSeq2003bs:*
10: GeneSeq2003cs:*
11: GeneSeq2003ds:*
12: GeneSeq2004as:*
13: GeneSeq2004bs:*
14: GeneSeq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	563	99.8	564	13	ADT97901
2	551.4	97.8	29340	13	ABD33482
3	551.4	97.8	57501	14	AD213482
4	38.2	6.8	2161	8	ABX34807
5	38.2	6.8	2777	11	ADM01893
6	37.2	6.6	712	6	ABK34393
7	37.2	6.6	1957	4	ABK42424
8	37.2	6.6	1957	4	AA107453
9	37.2	6.6	1957	9	ADB60580
10	36.2	6.4	138203	14	AD251725
11	36	6.4	546	13	ADQ56739
12	35.8	6.3	704	2	AAQ94044
13	35.8	6.3	1263	6	ABO90455
14	35.6	6.3	568	5	ABV51994
15	35.4	6.3	1785	6	AA159511
16	35.2	6.2	954	14	ACL70503
17	35.2	6.2	4050	14	ACL64141
18	35	6.2	550	9	ACH25075
19	35	6.2	1338	6	ABL58452

20	35	6.2	1338	12	ADP48674	Adp48674 Human Pel
21	35	6.2	1797	4	ABK42425	ABK42425 Genomic s
22	35	6.2	1797	4	AA107455	AA107455 Human rep
23	35	6.2	1797	9	ADB60581	ADB60581 Connectiv
24	34.8	6.2	792	13	ABD62531	ABD62531 Colton CD
25	34.8	6.2	76180	13	ABD33385	ABD33385 Human can
26	34.6	6.1	40742	4	AAK68089	AAK68089 Human imm
27	34.6	6.1	40742	4	AAK79886	AAK79886 Human imm
28	34.6	6.1	220890	6	AA138337	AA138337 Complemen
29	34.6	6.1	220895	6	ABK84798	ABK84798 Human CDN
30	34.6	6.1	220895	13	ADK52737	ADK52737 Drug Ther
31	34.4	6.1	5051	13	ADQ39151	ADQ39151 Human SNP
32	34.4	6.1	5232	13	ADQ39152	ADQ39152 Human SNP
33	34.4	6.1	92638	6	ABO88096	ABO88096 Human ost
34	34.2	6.1	756	2	AAK85016	AAK85016 Human sec
35	34.2	6.1	756	8	ACD18942	ACD18942 Novel hum
36	34.2	6.1	756	12	ADG78333	ADG78333 Human sec
37	34.2	6.1	756	12	ADN60624	ADN60624 Human sec
38	34.2	6.1	349980	6	ABQ81844	ABQ81844 Bifidobac
39	34	6.0	593	12	ACH71544	ACH71544 Human gen
40	34	6.0	1028	10	ADD84538	ADD84538 12P1P1 v
41	34	6.0	1028	10	ADD84540	ADD84540 12P1P1 v
42	34	6.0	1028	12	ADM83831	ADM83831 Human can
43	34	6.0	1028	12	ADM83794	ADM83794 Human can
44	34	6.0	1028	12	ADM83796	ADM83796 Human can
45	34	6.0	1988	13	ADS56480	ADS56480 Bacterial

ALIGNMENTS

RESULT 1	ADT97901	ADT97901 standard; DNA, 564 BP.
ID	ADT97901	
XX	ADT97901;	
AC		
XX		
DT	27-JAN-2005	(first entry)
XX		
DE	Human fibroblast growth factor (FGF)-3 promoter 5' proximal region.	
XX		
KW	Human; ds, fibroblast growth factor-3; FGF-3; promoter; SNP;	
KW	single nucleotide polymorphism; cancer; oesophageal cancer; 5'UTR;	
KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;	
KW	oesophageal squamous cell carcinoma.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	variation	replace(69,Y)
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT		/note= "This SNP is specifically claimed in claim 6"
XX		
PN	US2004219582-A1.	
XX		
PD	04-NOV-2004.	
XX		
PF	11-MAR-2004; 2004US-00798652.	
XX		
PR	17-MAR-2003; 2003US-0455689P.	
XX		
PA	(GUOY/) GUO Y.	
XX		
PI	Guo Y;	
XX		
DR	WPI; 2004-794435/78.	
XX		
FT	Novel isolated nucleic acid molecule having single nucleotide	
FT	polymorphism in upstream untranslated region of fibroblast growth factor-	
PT	3 gene, useful for assessing related susceptibility of mammal to cancer.	
XX		
PS	Claim 1; SEQ ID NO 1; 15pp; English.	

XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (AD97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to AD97901, a vector comprising
CC AD97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of AD97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC AD97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of AD97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of AD97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is the upstream untranslated region (UTR)
CC of fibroblast growth factor (FGF)-3 gene sequence.

XX Sequence 564 BP; 117 A; 160 C; 185 G; 101 T; 0 U; 1 Other;

XX Query Match 99.8%; Score 563; DB 13; Length 564;

XX Best Local Similarity 99.8%; Pred. No. 2,3e-142;

XX Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACCTCAGCGTGACTCACCCCATGTG 60
DB 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACCTCAGCGTGACTCACCCCATGTG 60
QY 61 GGTGAGGTGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGTGGTGAACGGA 120
DB 61 GGTGAGGTGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGTGGTGAACGGA 120
QY 121 GGGGCTTTGGCATGTGGGCAACAGGGGCTGCATTAAGCTTACTCATGTGCAATCGAGTCC 180
DB 121 GGGGCTTTGGCATGTGGGCAACAGGGGCTGCATTAAGCTTACTCATGTGCAATCGAGTCC 180
QY 181 CTGGTGCACAGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAAGTGTGTG 240
DB 181 CTGGTGCACAGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAAGTGTGTG 240
QY 241 GCCGGCCATGCCCCCAAGTTGGCACTCATGCTTTGGCAGGGTGGGGCTTCCAGTCC 300
DB 241 GCCGGCCATGCCCCCAAGTTGGCACTCATGCTTTGGCAGGGTGGGGCTTCCAGTCC 300
QY 301 ACAGGGTCCCATCAAGTACAGCCAGGTGGCTGAGAAAGGTCCCTGCAGTCAATGAAA 360
DB 301 ACAGGGTCCCATCAAGTACAGCCAGGTGGCTGAGAAAGGTCCCTGCAGTCAATGAAA 360
QY 361 CCAAGGAGGCTTTGGAAAACAATCTGAAGGCAATGCTTTGATTAGTGAAGAGGTGG 420
DB 361 CCAAGGAGGCTTTGGAAAACAATCTGAAGGCAATGCTTTGATTAGTGAAGAGGTGG 420
QY 421 GCGTGGGCTGGGGAAGGCCACAGGTCTGAGTCAAGGCCAGAGGCAAGAGTGGTCCCC 480
DB 421 GCGTGGGCTGGGGAAGGCCACAGGTCTGAGTCAAGGCCAGAGGCAAGAGTGGTCCCC 480
QY 481 AGCAGTCCCGCGGCTCTGGAGTGAAGTCTCTGGCCCACTGAGAAACGCTGTGAAG 540
DB 481 AGCAGTCCCGCGGCTCTGGAGTGAAGTCTCTGGCCCACTGAGAAACGCTGTGAAG 540
QY 541 AGGCAATGGCGTCTTTCCGACTTC 564
DB 541 AGGCAATGGCGTCTTTCCGACTTC 564

DB 541 AGGCAATGGCGTCTTTCCGACTTC 564

RESULT 2

ABD33482
ID ABD33482 standard; DNA; 29340 BP.

ABD33482;

18-NOV-2004 (first entry)

Human cancer-associated (CA) gene HD07-093.

Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;

de; cancer; cytostatic.

Homo sapiens.

WO2004058146-A2.

15-JUL-2004.

15-DEC-2003; 2003WO-US040081.

17-DEC-2002; 2002US-00322281.

(SAGRE-) SAGRES DISCOVERY INC.

Morris DW, Malandro MS;

WP1; 2004-499109/47.

Novel human cancer associated protein encoded within open reading frame
of cancer associated gene, useful as targets for diagnosing cancer.

Claim 16; SEQ ID NO 642; 182bp; English.

The invention relates to cancer-associated proteins (CAP) and the cancer-
associated (CA) nucleic acids encoding them. The invention also relates
to a method for treating cancers involving administering to a patient an
inhibitor of CAP, and a method of screening for anticancer activity in a
potential drug involving providing a cell that expresses a CA gene,
contacting a tissue sample derived from a cancer cell with an anticancer
drug candidate and monitoring the effect of the anticancer drug candidate
on expression of the CA gene. The CAP proteins are useful for detecting
cancer associated with expression of a CAP protein in a test cell sample
and for screening for a bioactive agent capable of modulating the
activity of a CAP protein. The CA nucleic acids are useful for diagnosing
cancer, involving determining the expression of a CA nucleic acid in a
tissue. This sequence represents a human CA gene of the invention. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 29340 BP; 6351 A; 8588 C; 8017 G; 6384 T; 0 U; 0 Other;

XX Query Match 97.8%; Score 551.4; DB 13; Length 29340;

XX Best Local Similarity 99.6%; Pred. No. 9.3e-139;

XX Matches 563; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACCTCAGCGTGACTCACCCCATGTG 60
DB 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACCTCAGCGTGACTCACCCCATGTG 60
QY 61 GGTGAGGTGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGTGGTGAACGGA 120
DB 61 GGTGAGGTGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGTGGTGAACGGA 120
QY 121 GGGGCTTTGGCATGTGGGCAACAGGGGCTGCATTAAGCTTACTCATGTGCAATCGAGTCC 180
DB 121 GGGGCTTTGGCATGTGGGCAACAGGGGCTGCATTAAGCTTACTCATGTGCAATCGAGTCC 180
QY 181 CTGGTGCACAGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAAGTGTGTG 240
DB 181 CTGGTGCACAGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAAGTGTGTG 240

Qy	199	AGCTGGAAGTATGACAAATTTTCCCATTAAGGAAAGTGTGTGGCCGGCATGCCCCCA	258
Db	225	AGTAGACAGGAGGCACTTCTTCTCCTTAATGCAACCCAGCAGCCAGACCGAAGCCGATGCG	166
Qy	259	CGTTGCACACTACTGTCCTTTGCAAGGTTGG	289
Db	165	CATGACAAAGGCGCCGGTAGATATAGCAGG	135

RESULT 6
ABK34393/c
ID ABK34393 standard; cDNA; 712 BP.

XX	Sequence	712 BP; 141 A; 159 C; 205 G; 207 T; 0 U; 0 Other;
+	Query Match	6.6%; Score 37.2; DB 6; Length 712;
+	Best Local Similarity	56.6%; Pred. No. 6.7;
+	Matches	69; Conservative 0; Mismatches 53; Indels 0; Gaps 0
Qy	251	CCCCCAAGTTGACACTCACTGCTTTGACAGGGTTGGGCTTCCAGTCACAGGGTCCC 310
Db	560	CCGCCCATCTGGCTCACTGCTCTGGCCCTTGAGAGGAAGGGGCGCAGAGGGCTTCCCTGCC 501
Qy	311	ATCCAGTACAGCCCGAGGTGGCTTCACAAAGTCTCCCTGCGAGTCATGAACCAAGGAGAG 370
Db	500	AGGGAACCTCCACACACAGAGATCTGCAGACGGGGTGGCTTGGCAGTCAAGTCCACGAAACA 441
Qy	371	CT 372
Db	440	CT 439

```
RESULT 7
ABK42424
ID ABK42424 standard; DNA; 1957 BP
...
```

PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232808P.
PR 08-SEP-2000; 2000US-0232808P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246177P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259676P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-565190/63.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder such
PT as cancer or rheumatoid arthritis.
XX
XX
XX Disclosure; SEQ ID NO 1311; 673pp; English.
PS
XX
XX The present invention relates to the isolation of novel human connective
XX tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
XX (cDNA and genomic) sequences encoding them. The sequences of the
XX invention are useful in the diagnosis, treatment, prevention and/or
XX prognosis of diseases associated with connective tissue(s), including
XX cancer. The polynucleotide sequences of the invention are also useful in
XX gene therapy. ABR42102-ABR43116 represent genomic sequences encoding the
XX novel human connective tissue related polypeptides. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 1957 BP; 515 A; 580 C; 441 G; 421 T; 0 U; 0 Other;
SQ
Query Match 6 6%; Score 37.2; DB 4; Length 1957;
Best Local Similarity 56.6%; Pred. No. 8; 53; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 53;
QY 251 CCCCCAAGTTCACACTGCTTGGAGGTTGGGCTTCAGTCACAGGATCCC 310
DB 165 CCGCCATGCTGCTCACTGCTTCTGCTCCCTGAGGAAAGGCGGCTTCGTGCCC 224
QY 311 ATCAAGTACCAAGCCAGGTGCTGCAAGAGTCCCTGCACTCATGAAACCAAGGAGG 370
DB 225 AGGAACTCCACACAGAGAGTCTGACAGCGGCTGCGCTGGCACTCATGCCAGAAACA 284
QY 371 CT 372
DB 285 CT 286

RESULT 8
 ID AL07453
 AC AL07453;
 DT 21-NOV-2001 (first entry)
 DE Human reproductive system related antigen DNA SEQ ID NO: 10141.
 KW Human; reproductive system related antigen; reproductive system disorder;
 cancer; gene therapy; db.
 OS Homo sapiens.
 XX MO200155320-A2.
 XX 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US001339.
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225265P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227100P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229309P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.

[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0	6.6%;	37.2;	DB 9;	1957;
08-DEC-2000; 2000US-0251990P.				
11-DEC-2000; 2000US-0254097P.				
05-JAN-2001; 2001US-0259678P.				
17-JAN-2001; 2001US-00764847.				
(HUMA-) HUMAN GENOME SCI INC.				
Rosen CA, Ruben SM, Barash SC;				
WPI; 2003-634869/60.				
P-PSDB; ADB59454.				
New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.				
Disclosure; SEQ ID NO 1311, 248bp; English.				
The invention describes an isolated nucleic acid molecule (1), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA).				
The polypeptide or polynucleotide is useful for preventing, treating, or ameliorating medical conditions in a mammal. The connective tissue polypeptides, polynucleotides and antibodies are particularly useful for treating, preventing and/or prognosing disorders of connective tissues (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus, scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or neoplasias (e.g. leukemia), neurodegenerative disorders (e.g. Alzheimer's disease, or Parkinson's disease), cardiovascular diseases (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass complications), autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.				
05-DEC-2000; 2000US-0251990P.				
11-DEC-2000; 2000US-0254097P.				
05-JAN-2001; 2001US-0259678P.				
17-JAN-2001; 2001US-00764847.				
(HUMA-) HUMAN GENOME SCI INC.				
Rosen CA, Ruben SM, Barash SC;				
WPI; 2003-634869/60.				
P-PSDB; ADB59454.				
New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.				
Disclosure; SEQ ID NO 1311, 248bp; English.				
The invention describes an isolated nucleic acid molecule (1), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA).				
The polypeptide or polynucleotide is useful for preventing, treating, or ameliorating medical conditions in a mammal. The connective tissue polypeptides, polynucleotides and antibodies are particularly useful for treating, preventing and/or prognosing disorders of connective tissues (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus, scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or neoplasias (e.g. leukemia), neurodegenerative disorders (e.g. Alzheimer's disease, or Parkinson's disease), cardiovascular diseases (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass complications), autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.				
05-DEC-2000; 2000US-0251990P.				
11-DEC-2000; 2000US-0254097P.				
05-JAN-2001; 2001US-0259678P.				
17-JAN-2001; 2001US-00764847.				
(HUMA-) HUMAN GENOME SCI INC.				
Rosen CA, Ruben SM, Barash SC;				
WPI; 2003-634869/60.				
P-PSDB; ADB59454.				
New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.				
Disclosure; SEQ ID NO 1311, 248bp; English.				
The invention describes an isolated nucleic acid molecule (1), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA).				
The polypeptide or polynucleotide is useful for preventing, treating, or ameliorating medical conditions in a mammal. The connective tissue polypeptides, polynucleotides and antibodies are particularly useful for treating, preventing and/or prognosing disorders of connective tissues (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus, scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or neoplasias (e.g. leukemia), neurodegenerative disorders (e.g. Alzheimer's disease, or Parkinson's disease), cardiovascular diseases (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass complications), autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.				
05-DEC-2000; 2000US-0251990P.				
11-DEC-2000; 2000US-0254097P.				
05-JAN-2001; 2001US-0259678P.				
17-JAN-2001; 2001US-00764847.				
(HUMA-) HUMAN GENOME SCI INC.				
Rosen CA, Ruben SM, Barash SC;				
WPI; 2003-634869/60.				
P-PSDB; ADB59454.				
New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.				
Disclosure; SEQ ID NO 1311, 248bp; English.				
The invention describes an isolated nucleic acid molecule (1), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA).				
The polypeptide or polynucleotide is useful for preventing, treating, or ameliorating medical conditions in a mammal. The connective tissue polypeptides, polynucleotides and antibodies are particularly useful for treating, preventing and/or prognosing disorders of connective tissues (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus, scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or neoplasias (e.g. leukemia), neurodegenerative disorders (e.g. Alzheimer's disease, or Parkinson's disease), cardiovascular diseases (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass complications), autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.				
05-DEC-2000; 2000US-0251990P.				
11-DEC-2000; 2000US-0254097P.				
05-JAN-2001; 2001US-0259678P.				
17-JAN-2001; 2001US-00764847.				
(HUMA-) HUMAN GENOME SCI INC.				
Rosen CA, Ruben SM, Barash SC;				
WPI; 2003-634869/60.				
P-PSDB; ADB59454.				
New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.				
Disclosure; SEQ ID NO 1311, 248bp; English.				
The invention describes an isolated nucleic acid molecule (1), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA).				
The polypeptide or polynucleotide is useful for preventing				

```
FT /tag= c
FT /product= "Regulator protein, FscRI, SEQ ID 10"
FT CDS complement(4377..7205)
FT /tag= d
FT /product= "Regulator protein, FscRII, SEQ ID 11"
FT complement(7210..10320)
FT /tag= e
FT /product= "Regulator protein, FscRIII, SEQ ID 12"
FT complement(10298..13315)
FT /tag= f
FT /product= "Regulator protein, FscRIV, SEQ ID 13"
FT 13522..14898
FT /tag= g
FT /product= "Glycosyltransferase, FscMI, SEQ ID 17"
FT 14953..16011
FT /tag= h
FT /product= "GDP-ketosugar aminotransferase, FscMII, SEQ ID
FT 18"
FT 16106..17287
FT /tag= i
FT /product= "Cytochrome P-450 monooxygenase, FscP, SEQ ID
FT 14"
FT 17334..17528
FT /tag= j
FT /product= "Ferredoxin protein, FscFE, SEQ ID 15"
FT 17556..18413
FT /tag= k
FT /product= "Thioesterase, FscTE, SEQ ID 16"
FT 18610..20781
FT /tag= l
FT /product= "ADC synthase, PabAB, SEQ ID 21"
FT 20927..26158
FT /tag= m
FT /product= "PKS FscA, SEQ ID 2"
FT 26333..27340
FT /tag= n
FT /product= "ABC transporter, FscTI, SEQ ID 8"
FT 27561..28280
FT /tag= o
FT /product= "ABC transporter, FscTII, SEQ ID 9"
FT complement(28983..60860)
FT /tag= p
FT /product= "PKS FscC, SEQ ID 3"
FT complement(60962..77587)
FT /tag= q
FT /product= "PKS FscB, SEQ ID 4"
FT complement(77983..84132)
FT /tag= r
FT /product= "PKS FscF, SEQ ID 5"
FT /product= "PKS FscF, SEQ ID 5"
FT complement(84170..107485)
FT /tag= s
FT /product= "PKS FscE, SEQ ID 6"
FT complement(107496..136148)
FT /tag= t
FT /product= "PKS FscD, SEQ ID 7"
FT complement(136558..137766)
FT /tag= u
FT /product= "GDP-mannose-4,6-dehydratase, FscMIII, SEQ ID
FT 19"
FT
XX US2005089884-A1.
XX
XX 28-APR-2005.
XX
XX 06-APR-2004; 2004US-00819386.
XX
XX 23-OCT-2003; 2003KR-00074035.
XX
XX (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX (UYSH-) UNIV SHANGHAI JIAOTONG.
XX
XX Lee SY, Deng Z, Chen S, Jeong KJ, Zhou X,
XX
```

```
DR WPI; 2005-322033/33.
DR P-PSDB; ADZ51726, ADZ51727, ADZ51728, ADZ51729, ADZ51730, ADZ51731,
DR ADZ51732, ADZ51733, ADZ51735, ADZ51736, ADZ51737, ADZ51738,
DR ADZ51739, ADZ51740, ADZ51741, ADZ51742, ADZ51743, ADZ51744, ADZ51745,
DR ADZ51746.
XX
XX New gene cluster for the biosynthesis of FR-008 polyketides, useful for
XX increasing the productivity of FR-008 polyketides or to produce new FR-
XX 008 polyketide variants.
XX
XX Claim 1; SEQ ID NO 1; 22pp; English.
XX
XX The present invention relates to a gene cluster (ADZ51725) for the
XX biosynthesis of FR-008 polyketides derived from Streptomyces sp. FR-008.
XX FR-008 polyketide, a heptane macrolide having aglycone containing 4-
XX aminoacetophenone, has antifungal activity and also high toxicity against
XX mosquito larvae, and so is highly useful in agricultural and medical
XX fields. The gene cluster comprises the following genes: fscA, fscC, fscB,
XX fscF, fscE and fscD, which code for modular polyketide synthase (PKS);
XX fscI and fscTII, which code for ABC transporter proteins; fscRI, fscRII,
XX fscRIII and fscRIV genes which code for regulator proteins; fscP which
XX codes for cytochrome P450 monooxygenase; fscFE which codes for ferredoxin
XX protein; fscTE which codes for thioesterase; fscMI which codes for
XX glycosyltransferase; fscMII which codes for GDP-mannose-4,6-dehydratase;
XX aminotransferase; fscMIII which codes for GDP-mannose-4,6-dehydratase;
XX fscA which codes for FAD-dependent monooxygenase; pabAB which codes for 4
XX -amino-4-deoxychorismate (ADC) synthase and pabC which codes for ADC
XX lyase. The genes of the gene cluster of the invention can be used to
XX develop recombinant microorganisms capable of producing FR-008
XX polyketides, also to increase the productivity of the existing FR-008
XX polyketides or to produce new FR-008 polyketide variants, by its
XX modification.
XX
XX Sequence 138203 BP; 18153 A; 48202 C; 55342 G; 16506 T; 0 U; 0 Other;
XX
XX Query Match 6.4%; Score 36.2; DB 14; Length 138203;
XX Best Local Similarity 55.0%; Pred. No. 53;
XX Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
XX
XX QY 42 GGTGACCTACCCCATGCTGGTGGAGGTGAGGAGGACCTCTGAGGCAGGCGCAGGCAGC 101
XX Db 109131 GGTGAGGACACACACAGCGGGGCGTGGTCCGGGCGCTCGGACCGCTGACACAGGCTGCG 109190
XX
XX QY 102 CGTCAGTGGGTGACGCGAGGGGCTCTTGCATGTGTGGCAGAGGGCTGCATPACAGCTTA 161
XX Db 109191 GGTGAGGGCGCGCGCGCGGAGCGCGCGCGGAGGGCGCGCGTCCGTACCGGGTTC 109250
XX
XX QY 162 CTCAGTGC 170
XX Db 109251 CCGGCGCAGC 109259
XX
XX RESULT 11
XX ADQ56739
XX ID ADQ56739 standard; DNA, 546 BP.
XX
XX AC ADQ56739;
XX
XX DT 21-OCT-2004 (first entry)
XX
XX DE Novel canine microarray-related DNA sequence SeqIDB0041.
XX
XX KW canine microarray; drug screening; toxicity assay;
XX KW environmental pollutant; cellular response; gene expression profile;
XX KW toxic response; liver necrosis; fatty liver disease;
XX KW protein adduct formation; hepatitis; dog; de.
XX
XX OS Canis familiaris.
XX
XX PN W02004063324-A2.
XX
XX 29-JUL-2004.
XX
```


Db 550 GGATCTTCAGGGGGCCCTCCACACGACGACACGACCGGC 512

RESULT 13

ABQ90455/c
ID ABQ90455 standard; DNA; 1263 BP.

AC ABQ90455;

DT 01-OCT-2002 (first entry)

DE M. capsulatus gene #440 for DNA array.

KW Micro array; gene; de; differential expression; gene expression.

OS Methylococcus capsulatus.

PN WO200255655-A2.

PD 18-JUL-2002.

PR 14-JAN-2002; 2002WO-NO000019.

PR 12-JAN-2001; 2001NO-00000235.

PR 12-JAN-2001; 2001NO-00000239.

PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.

PI (TIGR-) TIGR.

PI Birkeland NK, Eldhammer I, Jonassen I, Jensen HB, Lien T,

PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;

PI Salberg SL;

DR WPI; 2002-557818/59.

PT Novel DNA array useful for determining differential expression of
PT Methylococcus capsulatus genes, comprises polynucleotides or
PT oligonucleotides representative for a selective number of Methylococcus
PT capsulatus genes.

PS Claim 13; Page 257-258; 678bp; English.

CC The invention relates to a novel DNA array giving a representation of a
CC number of Methylococcus capsulatus genes. The method of the invention is
CC useful for determination of the differential expression of the genes of
CC M. capsulatus, and for studying gene expression on a genomic scale and in
CC gene expression assays of M. capsulatus genes. The sequences shown in
CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the
CC invention

SQ Sequence 1263 BP; 272 A; 385 C; 401 G; 205 T; 0 U; 0 Other;

Query Match 6.3%; Score 35.8; DB 6; Length 1263;

Best Local Similarity 59.2%; Pred. No. 19;

Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 425 GGGGCTGGGAGGAGCCAGCTGTGAGTACAGAGCCAGAGGAGAGAGTGTCTCCACAGA 484

DB 437 GGGAGTGGAGGAGCGCCCTGGCGGCGCCGCGAGTAAAGGAGGAGCGGCGGCGCA 378

QY 485 CTGCGCGCGGCTGTGAGTACAGTCTCTGAGCCAGCTGAGA 527

DB 377 CTTTTCGCGCGCGGCGGCGGAGAGGCTCTCGAGAGCGCCAGAGA 335

RESULT 14

ABV51994/c
ID ABV51994 standard; cDNA; 568 BP.

AC ABV51994;

DT 17-SEP-2002 (first entry)

XX Human prostate expression marker CDNA 51985.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PR 20-FEB-2001; 2001WO-US0005171.

PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Nonahan JE;

DR WPI; 2001-662795/76.

PS Claim 1; Page 10092; 11750bp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 568 BP; 110 A; 169 C; 154 G; 135 T; 0 U; 0 Other;

Query Match 6.3%; Score 35.6; DB 5; Length 568;

Best Local Similarity 51.5%; Pred. No. 17;

Matches 106; Conservative 0; Mismatches 99; Indels 1; Gaps 1;

QY 302 CAGGATCCATTCACAGTACAGCCAGGCTGACAGAAAGTCCCTCCAGTATGAAG 361

DB 298 CCGGGGAGACACATGCTCTCCACCGAGGTTCCTGTCATGCCCTTCCACATGAGAGA 239

QY 362 CAAGGAGAGCTTGGGAACCATCTGAAGGAGCATGTGATTGATTAGTAGAGGGTGG 421

DB 238 GAGGTAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 179

QY 422 GCTGGGCTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 481

DB 178 CTTGAGCATTTGGGCTCTCTTCTTCCCAAGCCCTTGAGCATGCTTGGGGGCA 119

QY 482 G-CAGTGGCGCGGCTTGGGAGTC 506

DB 118 GTCCCTGCACACCTGCTCTGTGAGC 93

RESULT 15

AAS19511
ID AAS19511 standard; DNA; 1785 BP.

XX

Job time : 473 secs

AC AAS19511;

XX 26-MAR-2002 (first entry)

XX Reference sequence for human MPL gene exons 9-10.

XX Human; single nucleotide polymorphism; SNP; MPL; chromosome 1p34;
KW myeloproliferative leukemia virus oncogene; haplotyping; genotyping;
KW congenital amegakaryocytic thrombocytopaenia; CAMT; de.

XX Homo sapiens.

XX Key Location/Qualifiers

FT exon 614..773

FT /tag= a

FT /number= 9

FT Intron 774..1027

FT /tag= b

FT exon /number= 9

FT 1028..1124

FT /tag= c

FT /number= 10

XX WO200179232-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US012301.

XX 14-APR-2000; 2000US-0197839P.

XX (GENA-) GENAISSANCE PHARM INC.

XX Chew A, Choi JY, Koshy B, Stephens JC;

XX WPI; 2002-055251/07.

DR P-PSDB; AAU09495.

XX Nucleotide polymorphisms in the human myeloproliferative leukemia virus
PT oncogene (MPL) gene, useful for studying the function of and expressing
PT MPL protein for use in screening drugs for treating diseases related to
PT MPL activity.

PS Claim 27; Fig 3; 85pp; English.

XX The present invention relates to novel single nucleotide polymorphisms
CC (SNPs) in the human myeloproliferative leukemia virus oncogene (MPL)
CC gene located on chromosome 1p34, and methods for haplotyping and/or
CC genotyping the MPL gene. The methods of the invention make use of allele-
CC specific oligonucleotides (ASOs) as probes and primers and/or primer-
CC extension oligonucleotides for detecting MPL gene polymorphisms. The
CC polymorphisms and screened compounds are useful for the treatment of
CC diseases associated with MPL activity, such as congenital amegakaryocytic
CC thrombocytopaenia (CAMT). The present sequence represents a reference
CC sequence for human MPL gene exons 9-10. Note: This sequence encodes for
CC only part of the MPL protein shown in AAU09495

XX Sequence 1785 BP; 317 A; 562 C; 500 G; 403 T; 0 U; 3 Other;

XX Query Match 6.3%; Score 35.4; DB 6; Length 1785;

XX Best Local Similarity 61.3%; Pred. No. 26;

Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 53 CCCATGTCGTGAGGTAGGAGCTCTCTAGGAGGCGGAGGCGGAGCCCTCAGGTGG 112

DB 913 CGCAGGTGTCCTCCAGTCCCGAGGCGGCGGAGGCGGCGGAGTACGAGGCTGCTGGA 972

OY 113 TGACGGCAGGGGTCTTGCCATGTCGGGCAAG 145

DB 973 TGAGGCGGGGCTCCGGCCCGGCTGGGCGAAG 1005

Search completed: March 25, 2006, 16:28:08

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:12:32 ; Search time 3244 Seconds
(without alignments)
8134.376 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtgcgtcttcgacatc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_ests1:
2: gb_ests2:
3: gb_ests3:
4: gb_hnc:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_ests7:
9: gb_gss1:
10: gb_gss2:
11: gb_gss3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128.4	22.8	603	AG133394	AG133394 Pan trogl
2	41	7.3	312	A1680295	A1680295 tw62a11.x
3	40.4	7.2	925	CNS0091P	AL053013 Drosophi1
4	40	7.1	925	CNS0091P	AL053013 Drosophi1
5	39.6	7.0	915	B0948345	B0948345 AGENCOURT
6	39.6	7.0	1101	CNS0178Y	BA108460 Drosophi1
7	38.8	6.9	321	A1289637	A1289637 qw29d03.x
8	38.2	6.8	637	CL542202	CL542202 OB_Ba006
9	38.2	6.8	867	BQ931784	BQ931784 AGENCOURT
10	38	6.7	393	A1970105	A1970105 wq89c02.x
11	37.8	6.7	481	AWS19131	AWS19131 ha50f09.x
12	37.8	6.7	523	A1818367	A1818367 w53h08.x
13	37.6	6.7	592	CF795053	CF795053 891143 MA
14	37.4	6.6	554	AA975697	AA975697 EST187806
15	37.2	6.6	294	AA983391	AA983391 cq56e01.s
16	37.2	6.6	360	BG330743	BG330743 na39e01.s
17	37.2	6.6	369	BUN94839	BUN94839 AGENCOURT
18	37.2	6.6	392	A1651148	A1651148 wa97h01.x
19	37.2	6.6	392	BES02665	BES02665 h220e10.x
20	37.2	6.6	395	BES02405	BES02405 h220e10.x
21	37.2	6.6	406	A1365427	A1365427 qz09a08.x
22	37.2	6.6	414	A1571898	A1571898 to20h12.x

C	23	37.2	6.6	417	3	BP424001	BP424001
	24	37.2	6.6	442	1	AW501328	AW501328
	25	37.2	6.6	443	1	AW015811	AW015811
	26	37.2	6.6	450	1	A1363365	A1363365
	27	37.2	6.6	452	1	A1123638	A1123638
	28	37.2	6.6	461	1	A1126856	A1126856
	29	37.2	6.6	477	1	AA733044	AA733044
	30	37.2	6.6	488	1	A1571759	A1571759
	31	37.2	6.6	493	2	BB328847	BB328847
	32	37.2	6.6	499	1	A1096775	A1096775
	33	37.2	6.6	506	2	BF196583	BF196583
	34	37.2	6.6	513	1	A1858617	A1858617
	35	37.2	6.6	527	1	A1697948	A1697948
	36	37.2	6.6	557	1	AA952452	AA952452
	37	37.2	6.6	557	1	AA952452	AA952452
	38	37.2	6.6	564	5	BO671953	BO671953
	39	37.2	6.6	569	3	BM719509	BM719509
	40	37.2	6.6	573	3	BM682828	BM682828
	41	37.2	6.6	585	1	AA706790	AA706790
	42	37.2	6.6	586	3	BP366808	BP366808
	43	37.2	6.6	589	5	BO787648	BO787648
	44	37.2	6.6	608	6	CA309122	CA309122
	45	37.2	6.6	610	3	BM987383	BM987383
				636	3	BM675573	BM675573

ALIGNMENTS

RESULT 1
AG133394
LOCUS AG133394 603 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-146A21.R, genomic survey sequence.
ACCESSION AG133394
VERSION AG133394.1 GI:16663072
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
BAC and sequences of library PTB
Unpublished
2 (bases 1 to 603)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Aao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.sc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 603
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-146A21.R"
/sex="male"
/cell_type="lymphoblast"
/clone_id="PTB Chimpanzee Male BAC library"

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN

Query Match	22.8%	Score 128.4;	DB 10;	Length 603;
Best Local Similarity	99.2%;	Pred. No. 7.7e-23;		
Matches 129;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

DY 1 GCAGCCCTGCCTCAGAAAACAGAAAGACCGACACTCACGGTACTCACCCCATGTG 60
DB 473 GCAGCCCTGCCTCAGAAAACAGAAAGACCGACACTCACGGTACTCACCCCATGTG 532

61 GCTCGAGGCTGAAGGAGCCCTTCTTGAGGCAGGGCCAGGGCAGTCGTCAAGTGGGTGACGGCA 120
| | | | |
Db 533 GCTGGAGGCGAGGAGGCCCTTCTGAGGCAGGGCCAGGGCAGTCGTCAAGTGGGTGACGGCA 592

Db 593 GGGGTCTTGC 602

[illegible]

ACCESSION	AI680295
VERSION	AI680295.1
KEYWORDS	GI:4890477
EST.	
KEYWORDS	

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 312)	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

CONTACT INFORMATION
Unpublished (257)
Contact: Robert Strausberg, Ph.D.
Email: cga@rs-mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.

CDNA Library Preparation Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length 614 Std Error: 0.00
Seq. Primer: 400P from Gibco
High quality sequence stop: 503.

```

source
1..312
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TMAG8:2264252"
/tissue_type="poorly-differentiated endometrial
adenocarcinoma, 2 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP U3"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.45 kb. Life Technologies catalog #:
11541-018"

```

```
Query Match      7.3%; Score 41; DB 1; Length 312;
Best Local Similarity 54.2%; Pred. No. 8.4;
Matches 83; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
```

Db
150 CCGCCATGCTGGCTCAAGTTCCTCCTGCCCTGGAGGGAAGGGGCGAAGGGCTTCCCCTCC 209

Oy 311 ATCCAGTACAGCCAGGTCGTGCAAGAAAGTCCCTGCAAGTATGATGAAGCAAGAGAG 370
 Db 210 AGGAACTCCACCAAGAGATCTGCAACAGGGGCTGGCCCTGCAGTCAAGTCCAGGAACA 269
 Oy 371 CTTGGAAGACACATCTGAAGGCGATGCTTTG 403
 Db 270 CTGACCATCCCTTAGAGAGAACACAGCCCTTG 302

RESULT 3	
CNS0091P/c	
LOCUS	
CNS0091P	925 bp
DNA	linear
GSS	03-JUN-1999

BAC191D16 of RPECI-98 library from *Drosophila melanogaster* (fruit fly) genomic survey sequence.
AL053013

KEYWORDS	SOURCE	ORGANISM
GSS.		
Drosophila	melanogaster (fruit fly)	
Drosophila	melanogaster	
Eukaryota	Metazoa: Arthropoda: Hexanoda: Insecta: Pterygota:	

REFERENCE
AUTHORS
1 (pages 1 to 925)
Genoscope.

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage
BP 101 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP *Drosophila*

FEATURES

Macdonald-Gibson, DNA Library and Preparation of Technology, Department of Aeronautics and Astronautics, Princeton University, Princeton, NJ 08542-3208, USA. E-mail: macdonald-gibson@princeton.edu

Dr. Jong's laboratory is in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_id="RPC1-98"
/note="end : TE13"

```

Query Match	7.2%	Score 40.4	DB 10	Length 925
Best Local Similarity	13.2%	Pred. No. 15		
Matches 40	Conservative 135	Mismatches 127	Indels 0	Gaps 0

Db

Oy
250 gcccccccaagttgcacactcactgccccttgacagggttggcgcttcacagtcacacaggctc 309
::: :::: :
924 sbcsccscscsbcsccsssmstssnsbsbcscscssbsbsssstssmssssbsbsssscscsssss 865

[illegible]

804 GAGASHSSSSACBSSSSSSCCSASCASASSSSSSSSSSRRCCGAGCGSSGASSSSSSSA 745

Db 744 SAGSVSSASSSSSCSSSVASSMSCSSSSSSASASSSSSSSSASACASCCTT 605

TITLE Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.
JOURNAL OMAP Project
COMMENT Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert length: 145000 Std Error: 0.00
Plate: 0066 row: B column: 18
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES
Location/Qualifiers

1..637
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0066B18"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: PACIBAC1, Site_1: HindIII, Site_2: HindIII"

ORIGIN

Query Match 6.8%; Score 38.2; DB 10; Length 637;
Best Local Similarity 63.7%; Pred. No. 53;
Matches 58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 411 GAGAGGCTGGGCTGGCTGGCGCAAGCCACCAAGCTGTGATCAAGCCAGCGAGGA 470
DB 462 GCGAGAGATGGGGCGGAGGCGGCGAGACGACCAAGCGCGCAACCCCGAGCGAGGA 521
QY 471 GCTGTCCCGACGACTGCCCGCGGCTGTGC 501
DB 522 CCAGGCGCCAGGAGCTCCCGCGCGCGCGC 552

RESULT 9
BO931784 867 bp mRNA linear EST 21-AUG-2002
LOCUS BO931784
DEFINITION AGENCOURT_8803536 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6426298
5', mRNA sequence.
ACCESSION BO931784
VERSION BO931784.1 GI:22346815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2609 row: h column: 11
High quality sequence atop: 627.
Location/Qualifiers
1..867

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6426298"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Query Match 6.8%; Score 38.2; DB 5; Length 867;
Best Local Similarity 48.8%; Pred. No. 57;
Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 79 TCTTGAGGACGGGCCAGGCGCAGGCTGACGCGAGGGCTTTCATGTGG 138
DB 70 TCCGGCCCCAGAGTGGTGCCACGCTCGGGGGGTCACTGGCGAGCCAGAGCCCTGGGGG 129
QY 139 GCACAGGGGCTGCATACAGCTTACCTACAGTCAATCGAGTCCCTGGTGCAGCCTTGGGA 198
DB 130 CCACCGAGGCCCACTCGCTGTGTCACCCGAGACATCGCTTGGGCCCGGCACTAC 189
QY 199 AGTCTGAGAGTACGACATGTTTCCATTAGAGAAAGTGTGCGCGGCATGCCCCCA 258
DB 190 AGTGAAGAGAGGAGCACTGTCTCTTGATGAGACCCAGGAGCCAGGAGCCAGATGGC 249
QY 259 CGTTCACACTCATCTGCTTTGCAAGGGTTGG 289
DB 250 CATGACAAAGGCGCGGTGATGATGACGAG 280

RESULT 10
A1970105 393 bp mRNA linear EST 20-OCT-2000
LOCUS A1970105
DEFINITION w989c02.x1 NCI_CGAP_Gc6 Homo sapiens cDNA clone IMAGE:2479202 3',
mRNA sequence.
ACCESSION A1970105
VERSION A1970105.1 GI:5766931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 488 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source
1..393
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="IMAGE:2479202"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_idb="NCI CGAP GC6"
/note="vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneds
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

```

ORIGIN

Query Match	6.7%;	Score 38;	DB 1;	Length 393;
Best Local Similarity	54.2%;	Pred. No. 55;		
Matches 77;	Conservative	0;	Mismatches 65;	Indels 0;
			Gaps	0;

[illegible]

RESULT 11					
AM519131					
LOCUS	AM519131	481 bp	mRNA	linear	EST 03-MAR-2000
DEFINITION	hns0109.x1 NCT_CGAP_Pant	Homo sapiens	cDNA	clone IMAGE:2877161	3'
	hnsa sequence.				

ACCESSION	AM519131	GI:7157213
VERSION	AM519131.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
1 (bases 1 to 481)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Cactariini;
Homnidae; Homo.

TITLE	JOURNAL
National Cancer Institute, Cancer Genome Anatomy Project (CGAP)	Unpublished (1997)
Tumor Gene Index	Unpublished (1997)

Contact: Robert Strauberg, Ph.D.
Email: cgabps-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGR clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
image.lnll.gov/image/html/115sourcec.shtml
Seq. primer: -40UP from Gibco
High quality sequence: 419.

FEATURES
Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2877161"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_1b="NCI_GCAP_Pan1"
/notes="Organ: pancreas; Vector: pCMV-Sport6; site_1: SalI; site_2: NotI. Cloned unidirectionally. Primer: oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

```

ORIGIN

Query Match	6.7%	Score 37.8;	DB 1;	Length 481;
Best Local Similarity	56.6%	Pred. No. 64;		
Matches 69;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0.

Qy	251	CCCCCAAGTTGACACTCACTGCTTTGAGGGTTGGGGCTTCAGTCAAGGGTCC	310
Db	157	CCCCCATCTGGCTTCATTTCTGTGCCCTGAGGGAAAGGGGCTTCCCGTCC	216
Qy	311	ATCCACGTACAGCCAGGTGGCTGCAGTCAATGAACCAAGGGAGG	370
Db	217	AGGGAATCTCCACACAGGAGTGTGCAGACGGGCTGGCTGGCAGTCAATGCCAGGACA	276
Qy	371	CT	372
Db	277	CT	278

RESULT 12

LOCUS	523 bp	mRNA	linear	EST 07-MAR-2000
DEFINITION	AI818367			
KEYWORDS	Wk59H08.x1 NCI_CGAP_Pan1 Homo sapiens	CDNA	clone	IMAGE:2419743 3'
DESCRIPTION	mRNA sequence.			

ACCESSION	AI818367	GI:5437446
VERSION	AI818367.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS

TITLE	National Cancer Institute, Cancer genome Anatomy project
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 1852 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence. stop: 428.

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2419743"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1: Salt; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

```

ORIGIN

Query Match	6.7%	Score 37.8;	DB 1;	Length 523;
Best Local Similarity	56.6%	Pred. No. 65;		
Matches 69;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0

QY 251 CCCCCCAAGTTGACACACACATGCTCTTTCAGGGTTGGGGCTTCACAGTCAAGGGTCCC 310

Db 148 CCGCCCAAGCTGGCTCAATGTTCTCTGCCCCGTGAGAGGAAGGGGCGAAGGGCTTCGTCTCCC 207

QY 311 ATCCACGTACCAAGCCCAAGTGGCTGCAGAGAGTCCCTCGCAGTCATGAAACCAAGGAGG 370

Db 208 AAGCAACTCCACCAACAGAGAGTCTTCAACAGGGCTGGCCCTGGCAGTCAATCCAGGAACA 267

QY 371 CT 372
||
Db 268 CT 269

RESULT 13
LOCUS CF795053/c 592 bp mRNA linear EST 21-OCT-2003
DEFINITION 891143 MARC 4P1G Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF795053
VERSION CF795053.1 GI:37799626
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
Sus.

REFERENCE 1 (bases 1 to 592)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.B. and Keeler,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TM8015 row: K column: 23
Seq primer: TAGAAGCAGCTGAGG.
Location/Qualifiers
1..592
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/ciseq_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN

Query Match 6.7%; Score 37.6; DB 6; Length 592;
Best Local Similarity 50.6%; Pred. No. 76;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 353 TCATGAAACCAAGGAGCTTGGAAACACATCTGAAGGCGATGCTTTGATTTAGTGA 412
Db 234 TCAGGAAACCAAGGAGAGATGATTAATAATACGAGGGGGCTTGTGATGATGTCA 175

QY 413 GAGGCTGGGGCTGGGCTGGGCAAGGCCACAGGTGTGATGATGAGACCCAGAGCGAGAAC 472
Db 174 CCGGGCCCACTACTCTGGGAGGCTGTGAGGTGTGCGGCTACGCCCTGGCCAGCTGCT 115

QY 473 TGTCTCCAGCACTGCCCGCCCTCTGCGATGAGTCTCTGCGCCACTGAGAACAGC 532
Db 114 CCGTCCAGGGCTCGCGCTTCCGCTTCTACCTTCTGTGCTGTGACCAAGGCTCAGC 55

RESULT 14
LOCUS AM975697/c 554 bp mRNA linear EST 02-JUN-2000
DEFINITION EST387806 MAGE resequencees, MAGN Homo sapiens cDNA, mRNA sequence.
ACCESSION AM975697
VERSION AM975697.1 GI:8166915
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 554)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@tigr.org
Plate: 352
Seq primer: Forward.
Location/Qualifiers
1..554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequencees, MAGN"
/note="Vector: pBluescriptSKm"

ORIGIN

Query Match 6.6%; Score 37.4; DB 1; Length 554;
Best Local Similarity 55.9%; Pred. No. 84;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 43 GTGACTACCCCTGCTGCTGCTGAGGTGAGGAGCTCTGAGCGAGGCGGCGAGCC 102
Db 545 GTTCTTAAGCCCTTAATTTGCAAGGTGTGAACCCCTATGGGCGCTGGCCAAATCGCA 486

QY 103 GTCAAGTGGGTGACGCGAGGGGCTTGCATATGTTGGGCAACAGGGGCTGCATACGCTTAC 162
Db 485 AGAAGGGGGCTGAAGAAAGGCGACGCTTGTGGGCTTCACTGGGACTGGAGAGAGCTGCC 426

QY 163 TCACTGA 169
Db 425 TTCCTGA 419

RESULT 15
LOCUS AA983291 294 bp mRNA linear EST 23-JUL-1998
DEFINITION 0656601.81 NCI_CGAP_Kids Homo sapiens cDNA clone IMAGE:1590360 3',
mRNA sequence.
ACCESSION AA983291
VERSION AA983291.1 GI:3161816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 294)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bdrip/image/image.html

Insert Length: 890 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 273.
 Location/Qualifiers

FEATURES
 source 1..294
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1590360"
 /issue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP_Kids"
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAAATTCGGCGCGCAATATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
 Query Match 6.6%; Score 37.2; DB 1; Length 294;
 Best Local Similarity 56.6%; Pred. No. 84;
 Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	251	CCCCCAAGCTTGCACACTCACTGCTTGGGAGTTGGGCTTCAGTCACAGGGTCCC	310
Db	159	CCGCCATGCTGCTCAGTTCCTGCCCCCTGGAGGAGAGGGCCGAGGGCTTCCGTC	218
QY	311	ATCCACGTACACAGCCGAGTGGCTGCAGAGGTCCCTCGCAGTCATGAACCAAGGAGG	370
Db	219	AGGGAAGTCCACACAGAGAGTGTGCAGACGGGCTGGCTGGCAGTCAGTCCAGGAACA	278
QY	371	CT 372	
Db	279	CT 280	

Search completed: March 25, 2006, 18:16:22
 Job time : 3245 secs

Db 17468 CATGACAAAGCGCGGTGATGATGACGAGG 17438

RESULT 2

US-09-949-016-14107
Sequence 14107, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14107

LENGTH: 19319

TYPE: DNA

ORGANISM: Human

US-09-949-016-14107

Query Match

Best Local Similarity 56.6%; Score 37.2; DB 3; Length 19319;

Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCCAAGTTCACACTGCTTTCAGAGGTTGGGCTTCCAGTCACAGGGTCCC 310

Db 17937 CCGCCATGCTGCTCAGTTCCTGCGCCCTGAGAGGAAGGGCCGAGGGCTTCCGCTCCC 17996

QY 311 ATCCAGCTACACAGCCAGGCTGTCAGAAAGTTCCTCCAGTCATGAACCAAGGAGG 370

Db 17997 AGGGAAGTCCACACAGAGGAGTCTGAGACGGGCTGCTGAGTCAGTCCAGGAGACA 18056

QY 371 CT 372

Db 18057 CT 18058

RESULT 3

US-09-621-976-18033

Sequence 18033, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 18033

LENGTH: 474

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 16

OTHER INFORMATION: n=a, g, c or t

US-09-621-976-18033

Query Match

Best Local Similarity 6.5%; Score 36.6; DB 3; Length 474;

Matches 51; Conservative 148; Mismatches 173; Indels 0; Gaps 0;

QY 37 CTCACGTGACTACACCCCATGCTGAGAGTGAAGGACCTCTCTGAGCAGGCCACAG 96

Db 4 CTCCKSCYSGNCCYCCMARWRMSYKGGSSSSCYKSCMCMCMRYCSGSYK 63

QY 97 GCAGCCGTGAGTGGTACGCGCAGGGCTCTTCCATGTTGGGACACAGGGCTGATACA 156

Db 64 TTTTMMAMMTTCTTGGMARRRMSGGGKTTTMMCKSKRTYSCKMGRMKGYRMY 123

QY 157 GCTTACGTGATGACATGCTGCTGTCGACGCTTGAAGTCTGAAGTGAAGCAAT 216

Db 124 CYKAGCYTMMKRWYSSCCMYTTKGGSMWTTTMMRRRKYSKRWTKKKKTTWMA 183

QY 217 GTTCCCATTAAGAAAGTGTGTGCGGCATGCCCCCAAGTTCAGTCACTGCTGC 276

Db 184 AMCYTTRMSYMMMRRAAAKTYYYCMMSTKMCMACCMMCRARASCMSRSY 243

QY 277 TTTCGAGGTTGGGCTTCCAGTACAGGTTCCATCCAGTACAGCCAGTGGCTGC 336

Db 244 TTMCCYTYMMYKGGRTMMRGMMKRMWYMKKSKSMKSCMWRMRMARTTYTWA 303

QY 337 AGAAGTCCCTGCGCATGTAACCAAGGAGGCTTGGAAACCACTGAAGGCAAT 396

Db 304 WYTTTYRMCYWRKTTTCMMWYSRWGSMWTARAGAMWCTWYMAARKKXTW 363

QY 397 GCGTTGATTA 408

Db 364 WMAARGGMTWA 375

RESULT 4

US-09-902-540-6966/c

Sequence 6966, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

PRIOR FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 6966

LENGTH: 954

TYPE: DNA

ORGANISM: Myxococcus xanthus

US-09-902-540-6966

Query Match

Best Local Similarity 53.7%; Score 35.2; DB 3; Length 954;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 407 TAGTGAAGAGGTGGCTGCTGCGCAAGGCACAGCTGAGTACAGGCCAGAGCA 466

Db 464 TCGTGCGGTGCGCGCTTCGAGAGCCCAAGACACGTCCTGGCCGCGTCCGAGGC 405

QY 467 GGAAGTGTGCTCCAGCACTGCCCCGCGCTTGGAGTCAATCTCTCTGSCACTGAG 526

Db 404 CGGCTCCCTTGACGGGAATTCATACCGGAGCGGCATGCAATCCCTCGGCTTCTTG 345

QY 527 AACAGCTGTAGAGAG 542

Db 344 AGCTGCCAGTAGCAG 329

RESULT 5

US-09-902-540-604

Sequence 604, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

PRIOR FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 6966

LENGTH: 954

TYPE: DNA

ORGANISM: Myxococcus xanthus

US-09-902-540-6966


```
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15483
LENGTH: 6872
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(6872)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15483
```

```
Query Match
Best Local Similarity 54.4%; Pred. No. 16;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
QY 330 AGGCGATGCTTGAATTTAGTGAAGAGGCTGGGCTGGGCGAAGCCACAGCTCTG 449
DB 1635 AGGATGGGCTGAGTTAAGTGCAGAGGCTGCCCGCCAGAAAGCTCTTCTCCAA 1566
QY 450 AGTCAGAGCCAGAGGCGAGAACTGTCTCCAGCACTGCGCCGCTCTGCGATGAGT 509
DB 1565 AACCGAGGCAAGTCAAGCTTGCAGCTTGGCCACCTTACCTGCTCTTGAAGGTTGT 1506
QY 510 CCTCC 514
DB 1505 CTCCC 1501
```

RESULT 13

```
US-09-949-016-15980/C
Sequence 15980, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15980
LENGTH: 239527
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(239527)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980
```

```
Query Match 6.0%; Score 33.8; DB 3; Length 239527;
```

```
Best Local Similarity 48.7%; Pred. No. 52;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 336 CAGTGCTGCGAAGAGTCCCTCGAGTCATGAACCAAGGAGGCTTGGAAACCAT 385
DB 194917 CAGGAGACTTAACTAAGTCAAGCAGACAGACAGTCAACCATATGCTGAGCCCAAC 194858
QY 386 CTGAAGGGCATGCTTTATTTAGTGAAGGCTGGGCTGGGCTGGGCAAGCCCAAG 445
DB 194857 AGACACTCAACACCAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 194798
QY 446 TCTGAGTCAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 505
DB 194797 CTCGACTGGAGAGAGATGGCTGGGAGCTTGAACCAAGGAGTCCCTGAGACCTG 194738
QY 506 CAGTCTCC 514
DB 194737 CAGTCTTC 194729
```

RESULT 14

```
US-09-621-976-17202/C
Sequence 17202, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTS and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 17202
LENGTH: 364
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-17202
```

```
Query Match
Best Local Similarity 10.9%; Pred. No. 7.1; Length 364;
Matches 30; Conservative 126; Mismatches 120; Indels 0; Gaps 0;
```

```
QY 9 GCTTCAGAAAACAGAGAGCGACACATTCACGCTGACTGACCCCATGCTGAGG 68
DB 292 SCWYSRKGSMSMTGSRMSMTGSRMSMTGSRMSMTGSRMSMTGSRMSMTGSR 233
QY 69 TGAAGGAGCCCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 128
DB 232 KKGSTSSRYASTSGSSSKYMTCKRSKRCRYATVYVSCMMKWKYCMMSATYSG 173
QY 129 GCCATGTGGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 188
DB 172 RMYVYSCMSRYSCTYSRSGCSTGKWKGCYRYMYRGRMYRGRMYRGRMYRGR 113
QY 189 AGCTCTGGAAGTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 248
DB 112 WCRSTKRYRKTCAWMGAKGTWCMRMSGTGAASYRMYRMYRMYRMYRMYRMY 53
QY 249 TGCCCCCAACGTTGACACATCACTGCTTTGACAG 284
DB 52 SSGSGSWKSCGSGSYCRSYSSMGKSRCTCCMG 17
```

RESULT 15

```
US-09-949-016-94858
Sequence 94858, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 94858
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-94858

```

```

Query Match      6.0%; Score 33.6; DB 3; Length 601;
Best Local Similarity 50.6%; Pred. No. 8.4;
Matches 78; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

```

```

QY      346 CTGGCAGTCATGAAACCAAGGAGGCTTGGGAAACCATCTGAAAGGCGATGAGCTTTGAT 405
Db      149 CTGGTAAGCCAGGGCCACAGGGTGTGAGAGACGCTCAAGACATGGGGCGGGCTGCGG 208
QY      406 TTAGTGAGAGGCTGGGGCTGGGCTGGGCAAGGCCACAGGTCYTAGTCAGAGCCAGAGGC 465
Db      209 TGAGTAGCAGGGATGGGCTGGGCGAGGCTGGGGGACCCCTGGGGGCTCTCCATGAGACAAGA 268
QY      466 AGGAAGCTGTGCTCCCGACACTGCCCCGCGCTCT 499
Db      269 AGGAACAGCATCTGAGAGACTGACACCCCTCT 302

```

Search completed: March 25, 2006, 15:02:01
 Job time : 202.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 16:12:32 ; Search time 3244 Seconds
(without alignment)
8134.376 Million cell updates/sec

Title: SEQ1-69C

Perfect score: 564

Sequence: 1 gcagccctgcctcagaaac.....agtcgccttcctgcactc 564

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_eest1:*
2: gb_eest2:*
3: gb_eest3:*
4: gb_hc:*
5: gb_eest4:*
6: gb_eest5:*
7: gb_eest6:*
8: gb_eest7:*
9: gb_gest1:*
10: gb_gest2:*
11: gb_gest3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130	23.0	603	10 AG133394	AG133394 Pan trogl
2	41	7.3	312	1 A1680295	A1680295 tw62a11.x
3	40.4	7.2	925	10 CNS0091P	AL053013 Drosophi1
4	40	7.1	925	10 CNS0091P	AL053013 Drosophi1
5	39.6	7.0	915	5 BQ948345	BQ948345 AGENCOURT
6	39.6	7.0	1101	10 CNS0178Y	AL108460 Drosophi1
7	38.8	6.9	321	1 A1289637	AL1289637 gw29d03.x
8	38.2	6.8	637	10 CL542202	CL542202 OB_Ba006
9	38.2	6.8	867	5 BQ931784	BQ931784 AGENCOURT
10	38	6.7	393	1 A1970105	A1970105 wg89c02.x
11	37.8	6.7	481	1 A1818367	A1818367 wa59f09.x
12	37.8	6.7	523	1 A1818367	A1818367 wa59f09.x
13	37.6	6.7	592	6 CF795053	CF795053 891143 MA
14	37.2	6.6	294	1 AA983291	AA983291 cg56e01.8
15	37.2	6.6	360	2 BG230743	BG230743 natf39e01.
16	37.2	6.6	369	5 BUI94839	BUI94839 AGENCOURT
17	37.2	6.6	392	2 A1651148	A1651148 wa97h01.x
18	37.2	6.6	392	2 BE502665	BE502665 h220e10.x
19	37.2	6.6	395	2 BE502665	BE502665 h220e10.x
20	37.2	6.6	406	1 A1365427	A1365427 qz09a08.x
21	37.2	6.6	414	1 A1571898	A1571898 to20h12.x
22	37.2	6.6	417	3 BP424001	BP424001 BP424001

23	37.2	6.6	442	1 AW501328	AW501328 UI-HF-BP0
24	37.2	6.6	443	1 AW015811	AW015811 UI-H-B10P
25	37.2	6.6	450	1 A1363365	A1363365 qy55e08.x
26	37.2	6.6	452	1 A1123638	A1123638 ool6a04.x
27	37.2	6.6	461	1 A1126856	A1126856 qb95a05.x
28	37.2	6.6	477	1 AA733044	AA733044 zg79a09.8
29	37.2	6.6	488	1 A1571759	A1571759 cn19d12.x
30	37.2	6.6	489	2 BE328847	BE328847 hv97g07.x
31	37.2	6.6	493	1 A1096775	A1096775 qb56e03.x
32	37.2	6.6	506	2 BF196583	BF196583 tm98b07.x
33	37.2	6.6	513	1 A1858617	A1858617 w140f03.x
34	37.2	6.6	527	1 A1697948	A1697948 w619b05.x
35	37.2	6.6	557	1 AW952452	AW952452 EST364507
36	37.2	6.6	564	5 BQ671953	BQ671953 AGENCOURT
37	37.2	6.6	569	3 BM719509	BM719509 UI-E-EJ1-
38	37.2	6.6	573	3 BM682828	BM682828 UI-E-EJ1-
39	37.2	6.6	585	1 AA706790	AA706790 zj30b03.8
40	37.2	6.6	586	3 BP366808	BP366808 BP366808
41	37.2	6.6	589	5 BQ787648	BQ787648 1m13908.x
42	37.2	6.6	608	6 CA309122	CA309122 UI-H-FM1-
43	37.2	6.6	610	3 BM987383	BM987383 UI-H-DP0-
44	37.2	6.6	636	3 BM675573	BM675573 UI-E-EJ1-
45	37.2	6.6	643	5 BU947551	BU947551 1053912.x

ALIGNMENTS

RESULT 1
AG133394
LOCUS 603 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-146A21.R, genomic survey sequence.
ACCESSION AG133394
VERSION AG133394.1 GI:1663072
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Torok,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of library PTB
TITLE Unpublished
JOURNAL 2 (bases 1 to 603)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Torok,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Ageo Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@ac.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 603
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-146A21.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC library"
ORIGIN

	Query Match Similarity	23.0%	Score 130;	DB 10;	Length 603;
	Best Local Similarity	100.0%	Pred. No.	2,9e-23;	
	Matches 130;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Dy	1 GCAGCCCTGCGTTCACAAAAACAAGAAGCAGCACACTCAAGTGACTCATCCCCATGTG	60			
Dd	473 GCAGGCCCTGCCTCAGAAAACAGAAAGACGACAGCACTCA CGGTGA CTCA CCCC CAT GTG	532			
Oy	61 GCTGGAGCGGAGGAGCGCTCTCTGAGGAGCGAGGCGACGCGTAGCGGTGA CGGCA	120			
Dd	533 GCTGGAGCGGAGGAGCGCTCTCTCTGAGGAGCGAGGCGACGCGTAGCGGTGA CGGCA	592			
Oy	121 GGCGTTCTTTC	130			
Dd	593 GGCGTTCTTGC	602			
RESULT 2	A1680295	312 bp mRNA linear EST 15-DEC-1999			
Locus	twe2all.xl NCI_CGAP_Ut3 Homo sapiens CDNA clone IMAGE:2264252 3'				
DEFINITION	mRNA sequence.				
VERSION	A1680295				
KEYWORDS	A1680295.1 GI:4890477				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Hemo. 1 (bases 1 to 312) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)				
JOURNAL COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nhl.gov Issue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CNLA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bdrp/image.html Insert Length: 614 Std Error: 0.00 Seq primer: -40UP from Glbo High quality sequence scop: 303.				
FEATURES	Location/Qualifiers				
Source	1..312				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:2264252"				
	/tissue_type="poorly-differentiated endometrial adenocarcinoma, 2 pooled tumors"				
	/lab_host="DH10B"				
	/clone_lhb="NCI_CGAP Ut3"				
	/note="Organ: uterus; Vector: pCMV-SPORT6, Site_1: SalI, Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.45 kb. Life Technologies catalog #: 11541-018"				
ORIGIN					
QY	Query Match	7.3%; Score 41; DB 1; Length 312;			
	Best Local Similarity	54.2%; Pred. No. 8.4;			
Matches	83; Conservative	0; Mismatches 70; Indels 0; Gaps 0;			
DY	251 CCCCCAAGCTTGACACACACTGCTTTGAGAGGGTGGGGCTTCAAGCACAGSGTCC	310			
Dd	150 CCGCCAAGCTGCTCAGTCTCTGCCCCCTGGAGGAGAGGGGSCAGAGGGCTTCCGATTCC	209			

QY	311	ATCAGCGACAGCGCCAGGTGGCTGTCAGAAAGTCCCTTGCGAGTATGAAACCAAGGAG	370
Db	210	AGGAACTCCACACAGAGACTCTGCAGACCGGCTGGCCCTTGCACTAGTCCAGAAACA	269
QY	371	CTTGCGAAACCACTCATGTAAGGCGACTTGG	403
Db	270	CTAGACATCCCTTATAGGAGGAGCACCGCCCTTG	302
RESULT 3			
CNS0091P/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			

Qy	490	CGCCGCTCTGTGAGCAGTCTCTCGGCACTTGAGAAACAGCTGTGAGAGGAGG 549
Db	684	SMCCCTCTBMSAARSSSSSSSSSSSMASSSASSSSSSSSSSSSSSSGSA 625
Qy	550	CG 551
Db	624	MS 623
RESULT 4		
LOCUS		
DEFINITION	CNS0091P	925 bp DNA linear GSS 03-JUN-1999
ACCESSION	CNS0091P	
VERSION	AL053013	
KEYWORDS	AT053013	
SOURCE	GSS.	
ORGANISM	Drosophila melanogaster (fruit fly)	
REFERENCE	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR19p16 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
AUTHORS	AL053013	
TITLE	AT053013	
JOURNAL	1 GI:4934461	
COMMENT	Drosophila melanogaster (fruit fly)	
	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
	1 (bases 1 to 925)	
	Genoscope.	
	Direct Submision	
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequeigenoscope.cns.fr - Web : www.genoscope.cns.fr)	
	Determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
FEATURES		
Source	location/Qualifiers	
	1..925	
	/organism="Drosophila melanogaster"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:7227"	
	/clone="BACR19D16"	
	/clone_1b="RPC1-98"	
	/note="end : TET3"	
ORIGIN		
Query Match	7.1%; Score 40; DB 10; Length 925;	
Best Local Similarity	16.8%; Pred. No. 19; Mismatches 112; Indels 0; Gaps 0;	
Matches	45; Conservative 111; Mismatch 112; Indels 0; Gaps 0;	
Qy	235	TGTGTGGCGGCATGCCCCCAAGTTCACACTGCTTGTGAGGTTGGGCTT 294
Db	554	TSSGGYVKGSSSGSBSCCSSSCSSSCSCBCCSCSSCSBYCCSSBSBSKCSST 613
Qy	295	CCAGTCACAGGTCCTCCATCCACATACAGCCAGTCAGTGCCTGACAAAGTCCCTGCAGTC 354
Db	614	SBSCSCCCSSKSVCTSCSSSSSSSSSSSTSSSTSSSTSKSSSGSSSSSSSYTTSK 673
Qy	355	ATGAACAACAGAGGAGCTTGGGAAACACATCTGAAGGACATGGCTTGATTGATTAAGTGA 414
Db	674	STASGSSSWAGCGSSGTSSTSSSSSSSTSSSVSSGSKSTBSGSSBSGSSSS 733
Qy	415	GGGTGGGCTGGGCTGGGACCAAGGCTGAGTCAAGACCAAGCAAGAGAGAGCTG 474

Db		734	SSTSBBCTSTSSSSSSSYSSSTCSCTCCSSYSSTSSSSTSMGSTRGSSSSSVG	793
Qy		475	GTCGCCAGACTGCCCGCGCCTTGCG	502
		:	: : : :	:
Db		794	TSSSDSTSTCCSCCYMCTCSTYBMB	821
 RESULT 5				
BQ948345				
LOCUS				
DEFINITION		BQ948345	915 bp	mRNA
		AGENCOURT_8803539 NIH_MGC_101 Homo sapiens	linear	EST 21-AUG-2000
ACCESSION		5', mRNA sequence.		
VERSION		BQ948345		
KEYWORDS		BQ948345.1 GI:22363823		
SOURCE		EST.		
ORGANISM		Homo sapiens (human)		
		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
		Homnidae; Homo.		
REFERENCE		1 (bases 1 to 915)		
AUTHORS		NIH-MGC http://ngc.nci.nih.gov/.		
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL		Unpublished (1999)		
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cga@db-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LNCM2609 row: j column: 11 High quality sequence start: 24 High quality sequence stop: 233. Location/Qualifiers 1..915 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6426346" /issue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_id="NIH MGC 101" /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGCGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
 FEATURES				
source				
 ORIGIN				
Query Match		7.0%; Score 39.6; DB 5; Length 915;		
Best Local Similarity		51.7%; Pred. No. 24;		
Matches		90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;		
Qy		79	TCCTGAAGGCAAGGGCCAGGGCAGCGGTGAGTGGTGAACGGAGGGGTTTCGCATGTGTG	138
Db		69	TCCGGCCCCAAGAGGTGTGTCACCAGTCTGGGGGGGTCACTGGCGAGCCAAGGCTTGGGGG	128
Qy		139	GCAAGAAGGGGTGCACTACAATTACTCACTAGTACATCATCAAGTCCCTGTGTGCCAGCTTGGGA	198
Db		129	CCACCGGAGGCCCACTCTGTGTGTCCACCCGGAGCATGTGTTGGGCCCGGGACCTCAC	188
Qy		199	AGTCTGAAGTAGCAATGTTTTCCCATTAAGAAAGTGTGTGGCGGCGCATGCC	252
Db		189	AGTAGCAAGAAGGACCTGTCTTCTCTTGATGTGCACCAAGACGCCAAGAACCC	242
 RESULT 6				

LOCUS	CNS0175Y	1101 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence Sp6 end of BAC BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108460				
VERSION	AL108460.1	GI:5628764			
KEYWORDS	GSS.				
ORGANISM	Drosophila melanogaster (fruit fly)				
SOURCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydrozoidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope.				
REFERENCE	Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
JOURNAL	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.				
FEATURES	Location/Qualifiers				
source	1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN37L08" /clone_11b="DrosBAC" /plasmid="pBelobAC11" /note="end : SP6"				
ORIGIN					
Query Match	7.0%;	Score 39.6;	DB 10;	Length 1101;	
Best Local Similarity	14.0%;	Pred. No. 25;			
Matches	43;	Conservative 153;	Mismatches 111;	Indels 1;	Gaps 1;
OY	238 GTGCGCGGCATGCCCCCAACGTTGCACACTACTGCTTGCAGGGTGGGCTTCCA	297			
Db	762 GSGSGGGGGCCAGCGVCGAACSSASCSSAMCGVSSGSCSAGSCCGVSSCGAVAS	821			
OY	298 GTCCACAGGTCCTCCACTGACGACGCCACCGACGTCGTGCAGAGGTCCTCGCAGTCA	357			
Db	822 SASVSWSVTAASVAVSASVAGSMASGAIVSSGCRSSVAVSVAAAVSSVSSSSSVAS	881			
OY	358 AAACCAAGGAGGAGCTTGGGAAACCAACCAVCTGACAGGCAATGCTTGAATTGATGAGAGG	417			
Db	882 AAAVSASSSSASASMAVAIAAAVAVSVASVSSSSSCSSSSASAVVAVVAASAS	941			
OY	418 TGGGCTGTGGCTGGGCAAGCCACAG-GTCTGAGTCAAGAGCCAGAGCGAGAGCTGCT	476			
Db	942 SVSSSSSSSVSTSSASVSVSAVMSAVVSSSSSSSVSVVAIAAIAAAAAAAAA	1001			
OY	477 CCCACGACACTGCCCGCCGCTCTGCGAGTCAAGTCTCTCGGCGCACTAGAAACGCTGT	536			
Db	1002 AAASSSSAVAAVVAVSSSSSSSASSSSSSVSVSSSCVSVSVSSSVSVAVAS	1061			
OY	537 AGAGAGGC 544				
Db	1062 ASASASVS 1069				
RESULT 7					
LOCUS	A1289637	321 bp	mRNA	linear	EST 21-DEC-1998
DEFINITION	qw29d03.x1 NCL_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:11992485 3', mRNA sequence.				
ACCESSION	A1289637				

VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AI289637.1	GI:3932817	Homo sapiens (human)	NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap.	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
EST.							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 321)							
NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap.						
Unpublished (1997)							
Tumor Gene Index							
Contact: Robert Strausberg, Ph.D.							
Email: cgaps-remail.nih.gov							
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.							
Emmert-Buck, M.D., Ph.D.							
CDNA Library Preparation: Life Technologies, Inc.							
CNA Library Arrayed by: Greg Lennon, Ph.D.							
DNA Sequencing by: Washington University Genome Sequencing Center							
clone distribution: NCI-CCAP clone distribution information can be							
found through the I.M.A.G.E. Consortium/BLM at:							
www-bio.llnl.gov/bdrip/image/image.html							
Insert length: 2780 Std Error: 0.00							
Seq primer: -40UP from Gibco							
High quality sequence stop: 290.							
Location/Qualifiers							
1..321							
/organism="Homo sapiens"							
/mol_type="mRNA"							
/db_xref="taxon:9606"							
/clone="IMAGE:1992485"							
/tissue_type="serous papillary carcinoma, high grade, 2							
pooled tumors"							
/lab_host="DH10B"							
/clone_1b="NCI CGAP UC4"							
/note="Organ: uterus; Vector: PCMV-SPORT6; Site 1: SalI;							
Site 2: NciI; Cloned unidirectionally. Primer: Oligo dt.							
Average insert size 1.48 Kb. Life Technologies catalog #:							
11542-016"							
ORIGIN							
Query Match	6.9%;	Score 38.8;	DB 1;	Length 321;			
Best Local Similarity	57.4%;	Pred. No. 32;					
Matches	70;	Conservative	0;	Mismatches	52;	Indels	0;
Gaps	0;						
251	CCCCCAAGTTGACACTACTGCTTGCAGAGGTTGGGCTTCCAGTACAGAGGTTCCC	310					
171	CCGCCATGCTGGCTCACTTCTCTCCCTCGAGGGAAGGGCCGAGGGGCTTCTGTC	230					
311	ATCCACGTACCAAGCCAGGCTGCTGCAAGAGTCCCTCGACATGATGAACCAAGGAGG	370					
231	AGGAACTCCACACAGAGAGTGTGACAGCGGCTGTGCTGACATGATCCAGGAACA	290					
371	CT	372					
291	CT	292					
RESULT 8							
CL542202							
LOCUS							
DEFINITION	CL542202	637 bp	DNA	linear	GSS 14-JUN-2004		
OB	Ba0066B18.f OB_Ba	Oryza brachyantha	genomic clone				
OB	Ba0066B18 5', genomic survey sequence.						
CL542202							
ACCESSION							
VERSION							
KEYWORDS	GS5.						
SOURCE	Oryza brachyantha						
ORGANISM	Oryza brachyantha						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;							
Ehharctoidae; Oryzeae; Oryza.							
1 (bases 1 to 637)							
Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,							

TITLE Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
JOURNAL OMAP Project
COMMENT Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 14500 Std Error: 0.00
Plate: 0066 row: B column: 18
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

source
1. .637
Location/Qualifiers
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0066B18"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_id="OB_Ba"
/note="Vector: pAG1BAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.8%; Score 38.2; DB 10; Length 637;
Best Local Similarity 63.7%; Pred. No. 53;
Matches 58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 411 GAGAGGCTGGGCTGGCTGGGCAAGCCACCAGCTGTGAGTCAGAGCCAGAGCGAGAA 470
DB 462 GCGAGGATGGGCGCGGAGCGCGCGAGAGCAACAGAGCGCGCAACCGCCAGCGAGAA 521

OY 471 GCTGTCTCCAGCACTGCGCGCGCGCTCTGC 501
DB 522 CAGGGCCAGCAGCTCTCCCGCGCGCGCGC 552

RESULT 9
BO931784 867 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8803536 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6426238
DEFINITION 5', mRNA sequence.
ACCESSION BO931784
VERSION BO931784.1 GI:22346815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM2609 row: h column: 11
High quality sequence stop: 627.

FEATURES

source
1. .867
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6426238"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Query Match 6.8%; Score 38.2; DB 5; Length 867;
Best Local Similarity 48.8%; Pred. No. 56;
Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

OY 79 TCTGAGGCGAGGCGCAGGCGCAGCTGAGTGAGCGAGGCGCTTGGCTGGTG 138
DB 70 TCCGCCCCAGAGGTGTGCCCCAGCTCGGCGGCTCACTGCGGCGCAAGGCGCTGGG 129
OY 139 GCACAGGGGCTGCATACAGCTTACTCAGTACATGAGTCCCTGGTGGCCAGCTCTGGA 198
DB 130 CCACCGAGGCCCACTCCCTGGTGTCCACCCGAGCATGCTTGGGCCCCGACATCAC 189
OY 199 AGTCTGAGTGAAGCAATGTTCCCATTAAGAAAGTGTGCGCGCCATGCCCCCAA 258
DB 190 AGTAGAGAGGAGGACATGTTCTCTTGAATGAGCACCCAGGACGCCAGAGCGATGCG 249
OY 259 CGTTGCACTCACTGCTTGGTGGAGGTTGG 289
DB 250 CATGACAAAGCGCGCGGTGATGAGACAG 280

RESULT 10
A1970105 393 bp mRNA linear EST 20-OCT-2000
LOCUS w699c02.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2479202 3',
DEFINITION mRNA sequence.
ACCESSION A1970105
VERSION A1970105.1 GI:5766931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 488 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES

source
1. .393
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"

QY 371 CT 372
DB 268 CT 269

RESULT 13
LOCUS CF795053/c 592 bp mRNA linear EST 21-OCT-2003
DEFINITION 891143 MARC 4P1G Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF795053
VERSION CF795053.1 GI:37799626
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae; Suidae;

REFERENCE
AUTHORS 1 (bases 1 to 592)
TITLE Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.E. and Keele,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: TMM8015 row: K column: 23
Seq primer: TAGAAGCAGCTCGAGG.
Location/Qualifiers
1..592
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/issue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 6.7%; Score 37.6; DB 6; Length 592;
Best Local Similarity 50.6%; Pred. No. 75;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 353 TCATGAAACCAAGGAGGCTTGGGAAACACATCTGAAGGCGATGGCTTGAATTAATGA 412
DB 234 TCAGGAACCAAGGAGGAGATGATATTAATAATACGAGGGGGGCTTGTGATATGCA 175
QY 413 GAGGCTGGGGCTGGGCTGGGAGGCCACGAGTCTGATGACAGCCAGAGGAGCAAGC 472
DB 174 CCGGCGCCAACTACCTCGGGAGGCTGTGAGTGTGGCTACGCCCTGGCCAGCTGTG 115
QY 473 TGTCTCCAGCACTGCCCGCCCTCTGCGATGAGTCTCTGCGCCACTGAAACAGC 532
DB 114 CCGTCCAGGGGCTCGGCGCTTCCCTCTTACACTTCTGTGCTCTGTGACACGAGGCTCAGC 55

RESULT 14
LOCUS AA983291 294 bp mRNA linear EST 23-JUN-1998
DEFINITION o656e01.e1 NCI_CGAP_K1d5 Homo sapiens cDNA clone IMAGE:1590360 3',
mRNA sequence.
ACCESSION AA983291
VERSION AA983291.1 GI:3161816
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 294)
TITLE NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMIL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 890 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 273.
Location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1590360"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP K1d5"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AACTGGAAGATTCGCGCCGCAATATTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 6.6%; Score 37.2; DB 1; Length 294;
Best Local Similarity 56.6%; Pred. No. 84;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCAAGCTTGACACTGCTCTTGTGAGGCTTCCAGTCAGCAGGCTCC 310
DB 159 CCGGCCATGCTGGTCACTGCTCTGCCCCCTGAGGAGGAGGCGCTTCCCGTCC 218
QY 311 ATCCACGTACACGCCAGGTGGCTGCAAGAGTCCCTCGAGTCATGAACCAAGGAGG 370
DB 219 AGGGAATCCACACAGAGAGTCTGAGACGGGCTGGCAGTCAGTCCAGCAACA 278
QY 371 CT 372
DB 279 CT 280

RESULT 15
LOCUS BG230743 360 bp mRNA linear EST 09-FEB-2001
DEFINITION na139e01.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:4143433 3',
mRNA sequence.
ACCESSION BG230743
VERSION BG230743
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 360)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: M. Bento Soares, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 338.
 Location/Qualifiers

FEATURES
 source
 1..360
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4143433"
 /issue_type="lymphocyte"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Soares NPBC"
 /note="Organ: blood; Vector: pTT73D-Pac; Site 1: NotI;
 Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCCGCGGTTTCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pTT73 vector. Library
 is normalized; constructed in the laboratory of M. Bento
 Soares (University of Iowa)."

ORIGIN

Query Match 6.6%; Score 37.2; DB 2; Length 360;
 Best Local Similarity 56.6%; Pred. No. 87;
 Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	251	CCCCCAACGTTGCACATCACTGCTTTGCAAGGTTGGGGCTTCCAGTCACAGGGTCCC	310
Db	153	CGCCCATGCTGCTCAGTTCTGCTGAGGGAAGGGCCGAGGGGCTTCCGCTCCC	212
QY	311	ATCCAGTACAGCCCAAGTGGCTGCAGAAGGTCCTCGCAGTCATGAACCAAGGAGG	370
Db	213	AGGGAATCCACACAGAGAGTCTGCACGCGGCTGGCTGCGAGTCAGTCCAGGACA	272
QY	371	CT 372	
Db	273	CT 274	

Search completed: March 25, 2006, 18:16:21
 Job time : 3248 secs

Db 17468 CATGACAAAGCGCGGTGATGATGACGAG 17438

RESULT 2
US-09-949-016-14107
Sequence 14107, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14107
LENGTH: 19319
TYPE: DNA
ORGANISM: Human
US-09-949-016-14107

Query Match
Best Local Similarity 56.6%; Score 37.2; DB 3; Length 19319;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 251 CCCCCCAAGTTGACACTCAGCTTCTTGACAGGTTGGGCTTCAGTCACAGGGTCCC 310
Db 17937 CCGCCATGCTGCTCAGTTCTCTGCGCCCTGAGAGGAGGGCCGAGGGCTTCCGTCCC 17996

Qy 311 ATCCACGTACCAAGCCAGGTGCTGTCAGAAAGTCCCTGCAGTCATGAACCAAGGAGG 370
Db 17997 AGGGAAGTCCCAACAGAGAGTCTGACAGCGGGCTGCTGCAGTCAGTCCAGGAGACA 18056

Qy 371 CT 372
Db 18057 CT 18058

RESULT 3
US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match
Best Local Similarity 6.3%; Score 35.4; DB 3; Length 474;
Matches 51; Conservative 147; Mismatches 0; Gaps 0;

Qy 37 CTCACGGTACTCACCCCATGTGCTGGAGCCGAGGAGCCTCTGAGGACGAGGCCAG 96
Db 4 CTCCKYS:CYGSGNCCYCCMARKRMSYKGRMYKSRGSSCYKSCMCMBSRYCSGSYK 63

Qy 97 GAGCCGTCAGTGGTGCAGGAGGCTTGGCATGATGGAGGACAGGAGGCTGATCA 156
Db 64 TTTTAAWMMWTTTGGKMAARRMSGGGKTTTMMCSKKKTSCHMRKWKXKYYSRWYY 123

Qy 157 GCTTACTGATGACATGAGTCCCTGCTGTCAGCTTGAAGTCTGAAGTGAACAAT 216
Db 124 CYKGCATYMMRWYSSCCMMYTTGGGSMWTTTMMRRRKKSYKRWTKGKKKXTWMA 183

Qy 217 GTTCCCATTAAGAAAGTGTGGCCGCGCATGCCCCCAAGTTCGACACTACTGCC 276
Db 184 AMCYTTWRSYMMWRRAAAKTYYYCMMSKTMCMACCMMCCWRBARSCMRSY 243

Qy 277 TTTGAGAGGTGGGGCTTCAGTCACAGGGTCCCATCAGTCACGACCCAGGCTGTC 336
Db 244 TTMCTYYMMYKGRMYMMRGMMKSMYMMYKKSMMKGSCHMKRAMWARKTYYTWA 303

Qy 337 AGAAGTCCCTGCGAGTCATGAACCAAGGAGGCTTGGAAACCAATCTGAAGGAG 396
Db 304 WYTTYVRMCCYMRKTTYCMMWMSRWMSMTTARAGAMWCMYMYMAARKKXTW 363

Qy 397 GCGTTGATTA 408
Db 364 WAAARGGMTWA 375

RESULT 4
US-09-902-540-6966/c
Sequence 6966, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 6966
LENGTH: 954
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-6966

Query Match
Best Local Similarity 6.2%; Score 35.2; DB 3; Length 954;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 407 TAGTGAGAGGTGGGCTGGCTGGGCAAGCCACACGAGTCTGAGTCAGAGCCAGGCA 466
Db 464 TGTGCGGTGCGCGGCTTTCAGAGAGCCCAAGACAGTCTTGCGCGCGCTGGAGC 405

Qy 467 GGAAGCTGTCCCAAGCACTGCGCGCGCTTGTGAGATGACAGTCTCTGTGCGCACTGAG 526
Db 404 CGGCTCCCTTGACCGGATCCATACCGAGAGCCGACAGTCCCTCGGAGCTTCTTGC 345

Qy 527 AACAGCCTGTAGAGG 542
Db 344 AGCTGCCAGTAGCGAG 329

RESULT 5
US-09-902-540-604
Sequence 604, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 604
LENGTH: 4050
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(4050)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-604

Query Match 6.2%; Score 35.2; DB 3; Length 4050;
Best Local Similarity 53.7%; Pred. No. 5.5;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 407 TAGTGAAGAGGCTGGGCTGGGCTGGGCAAGGCCACAGCTGATGATGAGGCAAGGCA 466
Db 3587 TCGTCCGCGTCCGCGCTTCAGAGAGCCCCAGACAGCTCTTGGCGCTGCTGAGC 3646
Qy 467 GGAAGCTGATCCCGACGACTGCCGCGCTCTGAGATGAGTCTCTGGGCAAGCTGAG 526
Db 3647 CGGCTCCCTTGAACGGATCATATCCGAGAGGCGCATGAGTCCCTGGGCTTCTTGG 3706
Qy 527 AACAGCCTGTAGAGG 542
Db 3707 AGCTGCCAGTAGCGAG 3722

RESULT 6
US-09-843-905A-11
Sequence 11, Application US/09843905A
Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (513)..(513)
OTHER INFORMATION: unsure
US-09-843-905A-11

Query Match 6.2%; Score 35; DB 3; Length 1338;
Best Local Similarity 49.2%; Pred. No. 4.3;

Matches 92; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 364 AGGAGGCTTGGAAACCATCTGAAGGCAATGGCTTTGATTAGTGAAGGCTGGGCG 423
Db 794 AGAGCGGCTCTCTCATCACTGTTGGGCGACACTGCTGTGGCGACACCGCGGGCG 853
Qy 424 TGGCTGGGCAAGGCAACAGCTGTGATCAGAGCCAGAGGCGAGAGGCTGCTCCGAGC 483

Db 854 TGCTCGGGCTCCCACTGAAGCACTGAGAGGCCAGCGGACAGAGCAATGACAGCG 913
Qy 484 ACTGCGCGCGGCTCTGATGATGATGATCTCTGCGCACTGAGAACAGCCTGTAGAGG 543
Db 914 GCGCCAGTCCCGGCGGCTGACACTCTGGGCTTCCAGCGCCAGCGCGTGGCGCA 973
Qy 544 CAGTGC 550
Db 974 CAGCGCC 980

RESULT 7
US-09-949-016-13882
Sequence 13882, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13882
LENGTH: 42610
TYPE: DNA
ORGANISM: Human
US-09-949-016-13882

Query Match 6.2%; Score 35; DB 3; Length 42610;
Best Local Similarity 46.8%; Pred. No. 13;

Matches 110; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 253 CCCCAGCTTGACACTCACTGCTTTGACAGGCTTGGGCTTCCAGTACAGGGTCCAT 312
Db 13589 CCTGACAGCTTGACCACTGAGGCTGCTGTGATGATGAGGCCAGTCCAGAGAGCTG 13648
Qy 313 CCAGTACAGAGGCTGCTGCAAGAGTCTCTCCAGTACATGAACCAAGGAGCT 372
Db 13649 TCACCAAGCGGTAGACACTCTGACAGTGGGCTGCTTGTGACAGCTCTGTG 13708
Qy 373 TGGAAACCATCTGAAGGCAATGCTTTGATTAGTGAAGGCTGGGCTGGG 432
Db 13709 GGGCAGGCTGGGCTGAGGCCCAAGAAACATGTGACTGAAAGGAGAGGCGAG 13768
Qy 433 CAAAGCCAGCAAGTCTGATGATGAGGCCAGAGGCAAGGCTGCTCCAGACTG 487
Db 13769 CACTCCAGAGGGGTGACGACAGGACCGGCTGGGATGTGGGCACTGCTG 13823

RESULT 8
US-09-949-016-14499/c
Sequence 14499, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14499
; LENGTH: 11276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(11276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14499

Query Match 6.0%; Score 34; DB 3; Length 11276;
Best Local Similarity 51.3%; Pred. No. 17;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 346 CTCGCAGTCATGAAACCAAGGAGGCTTGGGAAACCACTCTGAAGGGCATGGCTTTGAT 405
DB 7710 CTGCTAAGCAGGCGCCACAGGGTGTGAGAGAGCGCTCAGACATGGGCGGGCTGCCAG 7651
QY 406 TTAGTAGAGGAGTGGGCTGGGCTGGGCAAGGCCACAGTCTGAGTCAGAGCCAGAGGC 465
DB 7650 TTAGTAGAGGAGGATGGGCTGGGCAAGGCCACAGTCTGAGTCAGAGCCAGAGGC 7591
QY 466 AGGAAGCTGTGCTCCAGCACTGCGCGCCCTCT 499
DB 7590 AGGAACAGCATCTGGAGACCTGACCACTCCCTCT 7557

RESULT 9
US-09-621-976-17564
; Sequence 17564, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17564
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17564

Query Match 6.0%; Score 33.8; DB 3; Length 435;
Best Local Similarity 51.7%; Pred. No. 6.6;
Matches 77; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 228 AGGAAAGTGTGGCGCGGCAATGCCCCCAAGCTTGACACTCTGCTTTGCAAGGTT 287
DB 166 AGGAAGTGTGCGCGCGCATTTTCCCGACGCGCAACTTGTGCGCTTGAGAGGGGA 225
QY 288 GGGGCTTCAGTCACAGGCTCCCATCCAGTCAAGCCAGGCTGCGAAGGTCCT 347
DB 226 GGAGGCGCCAGGAGTACTGCGACGCAACATCAAAACAGATGCAAAAGATTCAATTAT 285
QY 348 CGCAGTCATGAACCAAGGAGGCTTGCG 376
DB 286 CTCCATCCCTTAATCTGTGTCACTGTG 314

RESULT 10
US-09-949-016-15483/C
; Sequence 15483, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15483
; LENGTH: 6872
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(6872)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15483

Query Match 6.0%; Score 33.8; DB 3; Length 6872;
Best Local Similarity 54.4%; Pred. No. 16;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 390 AGGCATGCTTTATTATTAGTAGAGGCTGGGCTGGGAGGCCACCAAGTCTG 449
DB 1625 AGGATGGCTGATTAATTAAGTGAAGGCTGCCAGTCCCAAGAGGCTCTGCCAAG 1566
QY 450 AGTCAGACCAAGGACGAAAGCTGTCCCAAGACTGCCCCGCTCTTGATGCACT 509
DB 1565 AACCGGCAAGGTCAGCTTGGCGCACCTTACCCTGCTCTTGAAGGCTGTGT 1506
QY 510 CTCC 514
DB 1505 CTCC 1501

RESULT 11
US-09-949-016-15980/C
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980

Query Match 6.0%; Score 33.8; DB 3; Length 239527;
Best Local Similarity 48.7%; Pred. No. 52;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```

Query Match Similarity      6.0%; Score 33.6; DB 3; Length 601;
Best Local Similarity      50.6%; Pred. No. 8.3;
Matches 78; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

Qy      346 CTCGCAGTCATGTAACCAAGGAGGCGTTGGAAACCACTCTGAAGGCGCATGCTTGTAT 405
Db      149 CTGGTAAACCAAGGCGCACAGGCGTGTGTAAGAGACGGCTCAGACATGGGCGGCTGCCAG 208

Qy      406 TTAGTGAAGAGGTTGGGGCTGGGCTGGGCAAGGCCACACAGTCTGAAGTCAGAGCCAGAGCC 465
Db      209 TGAGTAGCAGAGGATGGGCTGGGCGAGGCTGGGAGACCCCTGGGGCTCCTCCATGAGACAGA 268

Qy      466 AGAAGCTGGTCCCGACGACACTGCCCGCGGCTCT 499
Db      269 AGGAACAGCATCTGGAAGACTGACCAACCCCTCTT 302

RESULT 13
US-09-949-016-12065
; Sequence 12065, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR APPLICATION NUMBER: 60/237,768

```

Query March	6.0%;	Score 33.6;	DB 3;	Length 108441;
Best Local Similarity	57.7%;	Pred. No. 46;		
Matches 60;	Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;

Qy	66	AGGAGAGGAGCCTCTCAGACAGGCGACGAGCCGTTAGTGTGACGCGACAGGCT	125
Db	81526	AGGGGTGAACCACTATGAGCCCTGGCCAATATGACAGAGAGGGGCTGGAGGAGGCA	81588
Qy	126	CTTGCCATGATGGGCACAGGAGGCTGACATACAGCTTACTCAGTGA	169
Db	81586	GCTGTCTTGGGGTACACTGGGACCTGGAGAGCGTGCCTTCCTCTTA	81629

```

RESULT 15
US-10-104-047-1958
: Sequence 1958, Application US/10104047
: Patent No. 6943241
:
GENERAL INFORMATION:
:
APPLICANT: HELIX RESEARCH INSTITUTE
:
TITLE OF INVENTION: No. 6943241el full length cDNA
:
FILE REFERENCE: HI-A0105
:
CURRENT APPLICATION NUMBER: US/10/104,047
:
CURRENT FILING DATE: 2002-03-25
:
PRIOR APPLICATION NUMBER:

```

; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1958
 ; LENGTH: 2508
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-1958

Query Match 5.9%; Score 33.4; DB 3; Length 2508;
 Best Local Similarity 48.7%; Pred. No. 15;
 Matches 91; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy	364	AGGAGGCTTGGAAACCACTCTGAAGGCGATGCTTTGATTAGTAGAGGGTGGGC	423
Db	797	AGGACGGCTCTCTCATCGACTGTGGGGCCACTGCTGTGGCGCACACCGCGGGGC	856
Qy	424	TGGGCTGGGCAAGGCCACAGTCTGATCAGAGCCAGAGCAAGAACTGGTCCCAAC	483
Db	857	TGCTGGGGCTCCACACTGAAGCACTGAGGCCCGAGCGCAAAATGCAGCGC	916
Qy	484	ACTGCCCGCGCTCTGCGATGCACTCTCTGGCCACCTGAGAAACAGCTGTAGAGAG	543
Db	917	GGCCCAAGTGCCTCCGTTGGGCTTACGACTCTGGCTTCCCAAGCCCGTGGCCGA	976
Qy	544	CAGTGGC	550
Db	977	CAGCGCC	983

Search completed: March 25, 2006, 15:01:59
 Job time : 202.5 secs


```
Db 227 GGGGCTTTCACATGAGGACAGAGGAGCTGACATGAGTTACTGATGCAATTCAGATCC 286
Qy 181 CTGGTGGCCAGCCTCTGGAAGTCTGGAAGAGCAATGTTCCATTAAGGAAGTGTG 240
Db 287 CTGGTGGCCAGCCTCTGGAAGTCTGGAAGAGCAATGTTCCATTAAGGAAGTGTG 346
Qy 241 GCGGCGCATGCCCCCAAGTTCAGACACTGACTGCTTTGACAGGGTGGGGCTTCCAGTC 300
Db 347 GCGGCGCATGCCCCCAAGTTCAGACACTGACTGCTTTGACAGGGTGGGGCTTCCAGTC 406
Qy 301 ACAAGGCTCCATTCACATGACAGAGCCAGGTGGTGCAGAAAGTCCCTCCAGTCATGAA 360
Db 407 ACAAGGCTCCATTCACATGACAGAGCCAGGTGGTGCAGAAAGTCCCTCCAGTCATGAA 466
Qy 361 CCAAGGGAAGGCTTGGGAACCAATCTGAAGGCAATGCTTTGATTTAGTGAAGAGGTG 420
Db 467 CCAAGGGAAGGCTTGGGAACCAATCTGAAGGCAATGCTTTGATTTAGTGAAGAGGTG 526
Qy 421 GCGTGGGCTGGGCAAGGCCACAGTCTGATCAGAGCCAGAGGCAAGGCTGTCCCC 480
Db 527 GCGTGGGCTGGGCAAGGCCACAGTCTGATCAGAGCCAGAGGCAAGGCTGTCCCC 586
Qy 481 AGCACTGCCCGCGCTCTGCGATGCAATCTCTGCGCACTGAGAAACAGCTGTAGAG 540
Db 587 AGCACTGCCCGCGCTCTGCGATGCAATCTCTGCGCACTGAGAAACAGCTGTAGAG 646
Qy 541 AGGCAGTGGCGCTTTTCGAGCTTC 564
Db 647 AGGCAGTGGCGCTTTTCGAGCTTC 670
```

```
RESULT 2
US-10-027-632-142127
; Sequence 142127, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142127
```

```
Query Match 100.0%; Score 564; DB 6; Length 700;
Best Local Similarity 99.8%; Pred. No. 2,8e-165;
Matches 563; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGCCCTGCTCAGAAAACAGAAAGACGACAGCACTGACGGTGACTGACCCCATGTG 60
Db 107 GGAAGCCCTGCTCAGAAAACAGAAAGACGACAGCACTGACGGTGACTGACCCCATGTG 166
```

```
Qy 61 GCTGAGGAGGAGGAGACCTCTGAGAGCAGGCGCAGAGGAGCGCTGACGGTGAAGCGCA 120
Db 167 GCTGAGGAGGAGGAGCCTCTGAGAGCAGGCGCAGAGGAGCGCTGACGGTGAAGCGCA 226
Qy 121 GGGGCTTGGCCATGATGGGGCAAGGGGGTGCATACAGTTTACTGATGACAAATCGAGTCC 180
Db 227 GGGGCTTGGCCATGATGGGGCAAGGGGGTGCATACAGTTTACTGATGACAAATCGAGTCC 286
Qy 181 CTGGTGGCCAGCCTCTGGAAGTCTGGAAGTGAACAATGTTCCATTAAGGAAGTGTG 240
Db 287 CTGGTGGCCAGCCTCTGGAAGTCTGGAAGTGAACAATGTTCCATTAAGGAAGTGTG 346
Qy 241 GCGGCGCATGCCCCCAAGTTCAGACACTGACTGCTTTGACAGGGTGGGGCTTCCAGTC 300
Db 347 GCGGCGCATGCCCCCAAGTTCAGACACTGACTGCTTTGACAGGGTGGGGCTTCCAGTC 406
Qy 301 ACAAGGCTCCATTCACATGACAGAGCCAGGTGGTGCAGAAAGTCCCTCCAGTCATGAA 360
Db 407 ACAAGGCTCCATTCACATGACAGAGCCAGGTGGTGCAGAAAGTCCCTCCAGTCATGAA 466
Qy 421 GCGTGGGCTGGGCAAGGCCACAGTCTGATCAGAGCCAGAGGCAAGGCTGTCCCC 480
Db 527 GCGTGGGCTGGGCAAGGCCACAGTCTGATCAGAGCCAGAGGCAAGGCTGTCCCC 586
Qy 481 AGCACTGCCCGCGCTCTGCGATGCAATCTCTGCGCACTGAGAAACAGCTGTAGAG 540
Db 587 AGCACTGCCCGCGCTCTGCGATGCAATCTCTGCGCACTGAGAAACAGCTGTAGAG 646
Qy 541 AGGCAGTGGCGCTTTTCGAGCTTC 564
Db 647 AGGCAGTGGCGCTTTTCGAGCTTC 670
```

```
RESULT 3
US-10-798-652-1
; Sequence 1, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yonjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: FGF-3 GENE AND METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/798,652
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: n is cytosine or thymine
US-10-798-652-1
```

```
Query Match 99.8%; Score 563; DB 8; Length 564;
Best Local Similarity 99.8%; Pred. No. 4.2e-165;
Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGCCCTGCTCAGAAAACAGAAAGACGACAGCACTGACGGTGACTGACCCCATGTG 60
Db 1 GGAAGCCCTGCTCAGAAAACAGAAAGACGACAGCACTGACGGTGACTGACCCCATGTG 60
Qy 61 GCTGAGGAGGAGGAGCCTCTGAGAGCAGGGCGCAGAGGCAAGCCGTGAGGTGGTGAAGCGCA 120
Db 61 GCTGAGGAGGAGGAGCCTCTGAGAGCAGGGCGCAGAGGCAAGCCGTGAGGTGGTGAAGCGCA 120
```


QY 121 GGGGCTTGGCCATGGTGGGACAGGGGCTGCATAGAGTTACTAGTGAACAATCGAGTCC 180
DB 121 GGGGCTTGGCCATGGTGGGACAGGGGCTGCATAGAGTTACTAGTGAACAATCGAGTCC 180
QY 181 CTGGTGGCCAGGCTCTGGAAAGTCTGGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTGTG 240
DB 181 CTGGTGGCCAGGCTCTGGAAAGTCTGGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTGTG 240
QY 241 GCGGGGCAATGGCCCCCAAGTTGCACTGAGCTTGTGGAGGGTGGGGGCTTTCAGATGC 300
DB 241 GCGGGGCAATGGCCCCCAAGTTGCACTGAGCTTGTGGAGGGTGGGGGCTTTCAGATGC 300
QY 301 ACAGGGTCCCATCCATCCAGTACAGCCAGGTGGCTGCAGAAAGTCCCTGCAGTCAATGAAA 360
DB 301 ACAGGGTCCCATCCATCCAGTACAGCCAGGTGGCTGCAGAAAGTCCCTGCAGTCAATGAAA 360
QY 361 CCAGAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGGCTTGAATTTAGTGAAGAGGTGG 420
DB 361 CCAGAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGGCTTGAATTTAGTGAAGAGGTGG 420
QY 421 GGGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAGGAGGAAAGCTGGTCCC 480
DB 421 GGGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAGGAGGAAAGCTGGTCCC 480
QY 481 AGCACTGCCCCGCGCTCTGCATGCAAGTCTCTGGCCACTTGAGAAACAGCCTGTAGAG 540
DB 481 AGCACTGCCCCGCGCTCTGCATGCAAGTCTCTGGCCACTTGAGAAACAGCCTGTAGAG 540
QY 541 AGGAGTGGCGCTCTTTCGGAAGTTC 564
DB 541 AGGAGTGGCGCTCTTTCGGAAGTTC 564

RESULT 4

US-10-322-281-642
; Sequence 642, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 642
; LENGTH: 29340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-642

Query Match 98.0%; Score 553; DB 7; Length 29340;
Best Local Similarity 99.8%; Pred. No. 9.9e-162;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGGCTTGGCTCAGAAAACAGAAAGAGCAGACATCTACGGGTACTCACCCCATGTG 60
DB 3566 GCAGGCTTGGCTCAGAAAACAGAAAGAGCAGACATCTACGGGTACTCACCCCATGTG 3625
QY 61 GCTGAGGCGAGGAGGCTCTCTGAGGCGAGGGCGAGGGAGCGTCAAGGTGGGTGACGGCA 120
DB 3626 GCTGAGGCGAGGAGGCTCTCTGAGGCGAGGGCGAGGGAGCGTCAAGGTGGGTGACGGCA 3685
QY 121 GGGGCTTGGCCATGGTGGGCAAGGGGCTGCATACGCTTACTAGTGAACAATCGAGTCC 180
DB 3686 GGGGCTTGGCCATGGTGGGCAAGGGGCTGCATACGCTTACTAGTGAACAATCGAGTCC 3745
QY 181 CTGGTGGCCAGGCTCTGGAAAGTCTGGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTGTG 240
DB 3746 CTGGTGGCCAGGCTCTGGAAAGTCTGGAAAGTGAACAATGTTTCCATTAAAGAAAGTGTGTG 3805

QY 241 GCGGGCAATGGCCCCCAAGTTGCACTGAGCTTGTGGAGGGTGGGGGCTTTCAGATGC 300
DB 3806 GCGGGCAATGGCCCCCAAGTTGCACTGAGCTTGTGGAGGGTGGGGGCTTTCAGATGC 3865
QY 301 ACAGGGTCCCATCCATCCAGTACAGCCAGGTGGCTGCAGAAAGTCCCTGCAGTCAATGAAA 360
DB 3866 ACAGGGTCCCATCCATCCAGTACAGCCAGGTGGCTGCAGAAAGTCCCTGCAGTCAATGAAA 3925
QY 361 CCAGAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGGCTTGAATTTAGTGAAGAGGTGG 420
DB 3926 CCAGAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGGCTTGAATTTAGTGAAGAGGTGG 3985
QY 421 GGGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAGGAGGAAAGCTGGTCCC 480
DB 3986 GGGTGGGCTGGGCAAGGCCACCAAGTCTGAGTCAAGGCCAGGAGGAAAGCTGGTCCC 4045
QY 481 A-GCACTGCCCCGCGCTCTGCATGCAAGTCTCTGGCCACTTGAGAAACAGCCTGTAGA 539
DB 4046 AGGCACTGCCCCGCGCTCTGCATGCAAGTCTCTGGCCACTTGAGAAACAGCCTGTAGA 4105
QY 540 GAGGAGTGGCGCTCTTTCGGAAGTTC 564
DB 4106 GAGGAGTGGCGCTCTTTCGGAAGTTC 4130

RESULT 5

US-10-027-632-142126
; Sequence 142126, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 142126
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142126

Query Match 98.0%; Score 552.6; DB 5; Length 855;
Best Local Similarity 99.6%; Pred. No. 7.9e-162;
Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGGCTTGGCTCAGAAAACAGAAAGAGCAGACATCTACGGGTACTCACCCCATGTG 60
DB 108 GCAGGCTTGGCTCAGAAAACAGAAAGAGCAGACATCTACGGGTACTCACCCCATGTG 167
QY 61 GCTGAGGCGAGGAGGCTCTCTGAGGCGAGGGCGAGGGAGCGTCAAGGTGGGTGACGGCA 120
DB 168 GCTGAGGCGAGGAGGCTCTCTGAGGCGAGGGCGAGGGAGCGTCAAGGTGGGTGACGGCA 227
QY 121 GGGGCTTGGCCATGGTGGGCAAGGGGCTGCATACGCTTACTAGTGAACAATCGAGTCC 180
DB 228 GGGGCTTGGCCATGGTGGGCAAGGGGCTGCATACGCTTACTAGTGAACAATCGAGTCC 287

```
OY 181 CTGTGCGCAGCCTCTGGAAGTCTGGAAGTGAAGTCCATTAAAGAAAGTGTG 240
DB 288 CTGTGCGCAGCCTCTGGAAGTCTGGAAGTGAAGTCCATTAAAGAAAGTGTG 347
OY 241 GCCGCGCATGCCCCCAAGCTTGGACACTCTGCTTTGGAGGGTGGGGCTTCCAGTC 300
DB 348 GCCGCGCATGCCCCCAAGCTTGGACACTCTGCTTTGGAGGGTGGGGCTTCCAGTC 407
OY 301 AAGAGTCCCATTCACGTACGTACAGCCAGGTGGCTGAGAAGTCCCTCGACATCATGAAA 360
DB 408 AAGAGTCCCATTCACGTACGTACAGCCAGGTGGCTGAGAAGTCCCTCGACATCATGAAA 467
OY 361 CCAAGGGAAGGCTTGGGAAACCAATCTGAAGGGCACTGCTTTGATTAGTAGAGGGTGG 420
DB 468 CCAAGGGAAGGCTTGGGAAACCAATCTGAAGGGCACTGCTTTGATTAGTAGAGGGTGG 527
OY 421 GGCTGGGCTGGGCAAGGCCACCAAGTCTGATGACAGCCAGGCAAGGAGCTGTGCCCC 480
DB 528 GGCTGGGCTGGGCAAGGCCACCAAGTCTGATGACAGCCAGGCAAGGAGCTGTGCCCC 587
OY 481 A-GCACTGCGCGCGCTCTGTGGAGTCACTCTCGGCGCACCTGAGAACAGCTGTAGA 539
DB 588 AAGCACTGCGCGCGCTCTGTGGAGTCACTCTCGGCGCACCTGAGAACAGCTGTAGA 647
OY 540 GAGGCAGTGGCGCTCTTTGGGACTTC 564
DB 648 GAGGCAGTGGCGCTCTTTGGGACTTC 672
```

```
RESULT 6
US-10-027-632-142126
; Sequence 142126, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142126
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142126
```

```
Query Match 98.0%; Score 552.6; DB 6; Length 855;
Best Local Similarity 99.6%; Pred. No. 7.9e-162;
Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
```

```
OY 1 GCAGCCTCTCTGAGAAACAGAGAGCCAGACACTACGAGTGAACACCCCATGTG 60
DB 108 GCAGCCTCTCTGAGAAACAGAGAGCCAGACACTACGAGTGAACACCCCATGTG 167
OY 61 GCTGAGGCGAGGAGGAGCTCTGAGGCAAGGCCAGGGCAAGCCGTCAAGTGGTGAAGGCA 120
```

```
DB 168 GCTGAGGAGGAGGAGGAGCTCTCTGAGGCAAGGCCAGGGCAGCCGCTCAGGTGAGCGGCA 227
OY 121 GGGGCTTGGCCATGGTGGGCAAGGAGGCTGATACAGCTTAAGTACATCAAGTCC 180
DB 228 GGGGCTTGGCCATGGTGGGCAAGGAGGCTGATACAGCTTAAGTACATCAAGTCC 287
OY 181 CTGTGCGCAGCCTCTGGAAGTCTGGAAGTGAAGTCCATTAAAGAAAGTGTG 240
DB 288 CTGTGCGCAGCCTCTGGAAGTCTGGAAGTGAAGTCCATTAAAGAAAGTGTG 347
OY 241 GCCGCGCATGCCCCCAAGCTTGGACACTCTGCTTTGGAGGGTGGGGCTTCCAGTC 300
DB 348 GCCGCGCATGCCCCCAAGCTTGGACACTCTGCTTTGGAGGGTGGGGCTTCCAGTC 407
OY 301 AAGAGTCCCATTCACGTACGTACAGCCAGGTGGCTGAGAAGTCCCTCGACATCATGAAA 360
DB 408 AAGAGTCCCATTCACGTACGTACAGCCAGGTGGCTGAGAAGTCCCTCGACATCATGAAA 467
OY 421 GGCTGGGCTGGGCAAGGCCACCAAGTCTGATGACAGCCAGGCAAGGAGCTGTGCCCC 480
DB 528 GGCTGGGCTGGGCAAGGCCACCAAGTCTGATGACAGCCAGGCAAGGAGCTGTGCCCC 587
OY 481 A-GCACTGCGCGCGCTCTGTGGAGTCACTCTCGGCGCACCTGAGAACAGCTGTAGA 539
DB 588 AAGCACTGCGCGCGCTCTGTGGAGTCACTCTCGGCGCACCTGAGAACAGCTGTAGA 647
OY 540 GAGGCAGTGGCGCTCTTTGGGACTTC 564
DB 648 GAGGCAGTGGCGCTCTTTGGGACTTC 672
```

```
RESULT 7
US-09-925-065A-734471
; Sequence 734471, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 734471
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-734471
```

```
Query Match 7.1%; Score 39.8; DB 4; Length 541;
Best Local Similarity 64.8%; Pred. No. 0.056;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
OY 4 GCCCTGCTTCAGAAACAGAGAGCCAGACACTACGAGTGAACACCCCATGTGCT 63
DB 442 GCCCTGCTTCAGAAACAGAGAGCCAGACACTACGAGTGAACACCCCATGTGCT 501
OY 64 GAGGCGAGGAGGAGCTCTCTGAGGCAAGGCCA 94
```


/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-146-731-10

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.24;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

QY 23 AAGACGACGACATCTACAGGTCAGTACACCCCATGTGCTGAGCGAGGAGGCTCTCT 82
DB 56 VALAABSVNMMPMWGDITNNMIDRFVRAHLHDHPDTPPLTTTISPQESDERKCNRYR 115
QY 83 GAGGACGGGCGGAGGCGGCTCAGTGGGTGACGCGAGGGCTTTGCCATGTGGGAC 142
DB 116 YRGLVQNDPAGISEEQCLYQIYIDELYGGLQRPSEDEKKLAEKASIGTYEDSTVAEV 175
QY 143 AGGGCTCATACAGCTTACATGACATGACAGTCCCTGGTCCGAGGCTCTGGAATC 202
DB 176 EKAAREEPEEBSAAEESNSDEDEVIPDIDVEVDVDELNOBVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAGCAATGTTTCCATTAGAAGAGTGGTGGCCGCGCATGCCCAACGTT 262
DB 236 VMLRKDXEABAIKHAALBEEKAMYSGRSRBRQREFRERKLRKRTISPSYARBDP 295
QY 263 GCACATCTACTGCTTTCAGAGGTTGGGCTTCCAGTCAAGGTCCTCCATCCAGTACCA 322
DB 296 TYDPYKSPSSSSSRSSRSRSPTPGREKITYFTISFGSDEBMAAAAAASGVTTGK 355
QY 323 GCCCAGTGGCTGCAAGAGTCCCTGCAGTCATGAACCAAGAGGAGGCTTGGAAACCA 382
DB 356 PPAPPOGPAFGRNASARRSSSSSSSABRTSSRSSSSSRSGGTYRSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGGGCTGGGAGGCGCAC 442
DB 416 RSRSRWSRSRSTRSRYSRGRHSGGSRDGRHRSRPPARRGGYPPRRRSRSHS 475
QY 443 AGGTCTGATCAGAGCGCAAGAGCAAGACTGTCTCCCGACACTGCCGCCCTCTGTG 502
DB 476 GDRYRGGRLRHSSSRSSWSLSPRSRLTTRSRSHSPSGRSRSGSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

RESULT 11
US-10-140-472-10

/ Sequence 10, Application US/10140472
/ Publication No. US2003013888A1
/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ APPLICANT: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME

/ FILE REFERENCE: P3330R1C168
/ CURRENT APPLICATION NUMBER: US/10/140,472
/ CURRENT FILING DATE: 2002-05-06
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-472-10

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.24;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

QY 23 AAGACGACGACATCTACAGGTCAGTACACCCCATGTGCTGAGCGAGGAGGCTCTCT 82
DB 56 VALAABSVNMMPMWGDITNNMIDRFVRAHLHDHPDTPPLTTTISPQESDERKCNRYR 115
QY 83 GAGGACGGGCGGAGGCGGCTCAGTGGGTGACGCGAGGGCTTTGCCATGTGGGAC 142
DB 116 YRGLVQNDPAGISEEQCLYQIYIDELYGGLQRPSEDEKKLAEKASIGTYEDSTVAEV 175
QY 143 AGGGCTCATACAGCTTACATGACATGACAGTCCCTGGTCCGAGGCTCTGGAATC 202
DB 176 EKAAREEPEEBSAAEESNSDEDEVIPDIDVEVDVDELNOBVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAGCAATGTTTCCATTAGAAGAGTGGTGGCCGCGCATGCCCAACGTT 262
DB 236 VMLRKDXEABAIKHAALBEEKAMYSGRSRBRQREFRERKLRKRTISPSYARBDP 295
QY 263 GCACATCTACTGCTTTCAGAGGTTGGGCTTCCAGTCAAGGTCCTCCATCCAGTACCA 322
DB 296 TYDPYKSPSSSSSRSSRSRSPTPGREKITYFTISFGSDEBMAAAAAASGVTTGK 355
QY 323 GCCCAGTGGCTGCAAGAGTCCCTGCAGTCATGAACCAAGAGGAGGCTTGGAAACCA 382
DB 356 PPAPPOGPAFGRNASARRSSSSSSSABRTSSRSSSSSRSGGTYRSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGGGCTGGGAGGCGCAC 442
DB 416 RSRSRWSRSRSTRSRYSRGRHSGGSRDGRHRSRPPARRGGYPPRRRSRSHS 475
QY 443 AGGTCTGATCAGAGCGCAAGAGCAAGACTGTCTCCCGACACTGCCGCCCTCTGTG 502
DB 476 GDRYRGGRLRHSSSRSSWSLSPRSRLTTRSRSHSPSGRSRSGSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

RESULT 12
US-10-141-761-10

/ Sequence 10, Application US/10141761
/ Publication No. US20030148432A1
/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William

